

Please type a plus sign (+) inside this box ☒
 Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

Approved for use through 10/31/2002. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

PTO/SB/05 (08-00)

UTILITY PATENT APPLICATION TRANSMITTAL

(Only for new nonprovisional applications under 37 CFR 1.53(b))

Attorney Docket No. 47675-14
First Inventor Markl, Isabel et al.
Title METHYLATION ALTERED DNA SEQUENCE
Express Mail Label No. EL695285386US

APPLICATION ELEMENTS

See MPEP chapter 600 concerning utility patent application contents.

1. ☐ Fee Transmittal Form (e.g., PTO/SB/17)
(Submit an original and a duplicate for fee processing)
2. ☐ Applicant claims small entity status.
See 37 CFR 1.27.
3. ☒ Specification [Total Pages 14]
(preferred arrangement set forth below)
 - Descriptive title of the invention
 - Cross Reference to Related Applications
 - Statement Regarding Fed sponsored R & D
 - Reference to sequence listing, a table, or a computer program listing appendix
 - Background of the Invention
 - Brief Summary of the Invention
 - Brief Description of the Drawings (if filed)
 - Detailed Description
 - Claim(s)
 - Abstract of the Disclosure
4. ☐ Drawing(s) (35 U.S.C. 113) [Total Sheets]
5. Oath or Declaration [Total Pages]
 - a. ☐ Newly executed (original or copy)
 - b. ☐ Copy from a prior application (37 CFR 1.63 (d))
(for continuation/divisional with Box 17 completed)
 - i. ☐ **DELETION OF INVENTOR(S)**
Signed statement attached deleting inventor(s) named in the prior application, see 37 CFR 1.63(d)(2) and 1.33(b).
6. ☐ Application Data Sheet. See 37 CFR 1.76

ADDRESS TO: Assistant Commissioner for Patents
Box Patent Application
Washington, DC 20231

7. ☐ CD-ROM or CD-R in duplicate, large table or Computer Program (Appendix)
8. Nucleotide and/or Amino Acid Sequence Submission (if applicable, all necessary)
 - a. ☒ Computer Readable Form (CRF)
 - b. Specification Sequence Listing on:
 - i. ☐ CD-ROM or CD-R (2 copies); or
 - ii. ☐ paper
 - c. ☒ Statements verifying identity of above copies

ACCOMPANYING APPLICATION PARTS

9. ☐ Assignment Papers (cover sheet & document(s))
10. ☐ 37 CFR 3.73(b) Statement (when there is an assignee) ☐ Power of Attorney
11. ☐ English Translation Document (if applicable)
12. ☐ Information Disclosure Statement (IDS)/PTO-1449 ☐ Copies of IDS Citations
13. ☐ Preliminary Amendment
14. ☒ Return Receipt Postcard (MPEP 503)
(Should be specifically itemized)
15. ☐ Certified Copy of Priority Document(s)
(if foreign priority is claimed)
16. ☐ Other:

17. If a CONTINUING APPLICATION, check appropriate box, and supply the requisite information below and in a preliminary amendment, or in an Application Data Sheet under 37 CFR 1.76:

☐ Continuation ☐ Divisional ☐ Continuation-in-part (CIP)

of prior application No. _____ / _____

Prior application information:

Examiner _____

Group / Art Unit: _____

For CONTINUATION OR DIVISIONAL /
Box 5b, is considered a part of the di
The incorporation can only be relied t



23504

PATENT & TRADEMARK OFFICE

from which an oath or declaration is supplied under
application and is hereby incorporated by reference.
submitted application parts.

☒ Customer Number or Bar Code Label

or ☐ Correspondence address below

Name	Davis Wright Tremaine LLP				
Address	2600 Century Square				
	1501 Fourth Avenue				
City	Seattle	State	WA	Zip Code	98101-1688
Country	US	Telephone	206-622-3150	Fax	2066287699

Name (Print/Type)	Barry L. Davison	Registration No. (Attorney/Agent)	P47,309
Signature	[Signature]		Date 10-27-2000

Burden Hour Statement: This form is estimated to take 0.2 hours to complete. Time will vary depending upon the needs of the individual case. Any comments on the amount of time you are required to complete this form should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, Washington, DC 20231. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Assistant Commissioner for Patents, Box Patent Application, Washington, DC 20231

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicants: Markl, Isabel et al.

Filing Date: October 27, 2000

Serial No.: to be assigned

For: METHYLATION ALTERED DNA SEQUENCES AS MARKERS ASSOCIATED
WITH HUMAN CANCER

Docket: 47675-15

Date: October 27, 2000

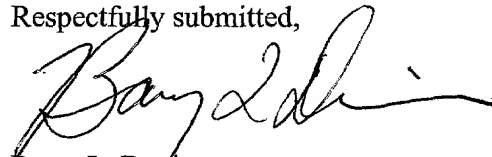
Assistant Commissioner for Patents
Box Patent Application
Washington, DC 20231

STATEMENT UNDER 37 C.F.R. §1.821

Sir:

I hereby state that the content of the paper and computer-readable copies of the Sequence Listing, submitted in accordance with 37 C.F.R. §1.821, are the same.

Respectfully submitted,



Barry L. Davison
Attorney for Applicants
Registration No. P47,309

Davis Wright Tremaine LLP
2600 Century Square
1501 Fourth Avenue
Seattle, WA 98101-1688
Tel: 206-628-7621
Fax: 206-628-7699

METHYLATION ALTERED DNA SEQUENCES AS MARKERS ASSOCIATED WITH HUMAN CANCER

5 Technical Field of the Invention

The present invention relates to novel human DNA sequences that exhibit altered methylation patterns (hypermethylation or hypomethylation) in cancer patients. These novel methylation-altered DNA sequences are useful as diagnostic, prognostic and therapeutic markers for human cancer.

10 Background of the Invention

The identification of early genetic changes in tumorigenesis is a primary focus in molecular cancer research. Characterization of the nature and pattern of cancer-associated genetic alterations will allow for early detection, diagnosis and treatment of cancer. Such genetic alterations in vertebrates fall generally into one of three categories: gain or loss of genetic material; mutation of genetic material; or methylation at cytosine residues in CpG dinucleotides within "CpG islands." Among these, DNA methylation is uniquely reversible, and changes in methylation state are known to affect gene expression (*e.g.*, transcriptional initiation of genes where CpG islands located at or near the promoter region) or genomic stability.

Methylation of CpG dinucleotides within CpG islands. DNA, in higher order eukaryotic organisms, is methylated only at cytosine residues located 5' to guanine residues in CpG dinucleotides. This covalent modification of the C-5 position of the cytosine base by the enzyme DNA (cytosine-5)-methyltransferase results in the formation of 5-methylcytosine (5-mCyt), and gives this base unique properties (*e.g.*, susceptibility to undergo spontaneous deamination). This enzymatic conversion is the only epigenetic modification of DNA known to exist in vertebrates, and is essential for normal embryonic development (Bird, A.P., *Cell* 70:5-8, 1992; Laird & Jaenisch, *Human Molecular Genetics* 3:1487-1495, 1994; Li et al., *Cell* 69:915-926, 1992).

The presence of 5-mCyt at CpG dinucleotides has resulted in the 5-fold depletion of this sequence in the genome during the course of vertebrate evolution (Schroeder & Gartler, *Proc. Nat. Acad. Sci. USA* 89:957-961, 1992), presumably due to spontaneous deamination of 5-mCyt to Thymidine. Certain areas of the genome, however, do not show such depletion, and are referred to as "CpG islands" (Bird, A.P., *Nature* 321:209-213, 1986; Gardiner-Garden & Frommer, *J. Mol. Biol.* 196:261-282, 1987). These CpG islands comprise only approximately 1% of the vertebrate genome, yet account for about 15% of the total number of genomic CpG dinucleotides (Antequera & Bird, *Proc. Nat. Acad. Sci. USA* 90:11995-11999, 1993). CpG islands contain the expected (*i.e.*, the non-evolutionarily depleted) frequency of

CpGs (with an Observed/Expected Ratio¹ >0.6), are GC-rich (with a GC Content² >0.5) and are typically between about 0.2 to about 1 kb in length.

Methylation within CpG islands affects gene expression. CpG islands are located upstream of many housekeeping and tissue-specific genes, but may also extend into gene coding regions (Cross & Bird, *Current Opinions in Genetics and Development* 5:309-314, 1995; Larsen et al., *Genomics* 13:1095-1107, 1992). The methylation of cytosines within CpG islands in somatic tissues is believed to affect gene expression. Methylation has been inversely correlated with gene activity and may lead to decreased gene expression by a variety of mechanisms including inhibition of transcription initiation (Bird, A.P., *Nature* 321:209-213, 1986; Delgado et al., *EMBO Journal* 17:2426-2435, 1998), disruption of local chromatin structure (Counts & Goodman, *Molecular Carcinogenesis* 11:185-188, 1994; Antequera et al., *Cell* 62:503-514, 1990), and recruitment of proteins that interact specifically with methylated sequences and thereby directly or indirectly prevent transcription factor binding (Bird, A.P., *Cell* 70:5-8, 1992; Counts & Goodman, *Molecular Carcinogenesis* 11:185-188, 1994; Cedar, H., *Cell* 53:3-4, 1988). Many studies have demonstrated the effect of methylation of CpG islands on gene expression (e.g., the *CDKN2A/p16* gene; Gonzalez-Zulueta et al., *Cancer Research* 55:4531-4535, 1995), but most CpG islands on autosomal genes remain unmethylated in the germline, and methylation of these islands is usually independent of gene expression. Tissue-specific genes are typically unmethylated in the respective target organs but are methylated in the germline and in non-expressing adult tissues, while CpG islands of constitutively expressed housekeeping genes are normally unmethylated in the germline and in somatic tissues.

Methylation within CpG islands affects the expression of genes involved in cancer. Data from a group of studies show the presence of altered methylation in cancer cells relative to non-cancerous cells. These studies show not only alteration of the overall genomic levels of DNA methylation, but also changes in the distribution of methyl groups. For example, abnormal methylation of CpG islands that are associated with tumor suppressor genes or oncogenes within a cell may cause altered gene expression. Such altered gene expression may provide a population of cells with a selective growth advantage and thereby result in selection of these cells to the detriment of the organism (*i.e.*, cancer).

Insufficient correlative data. Unfortunately, the mere knowledge of the basic existence of altered methylation of CpG dinucleotides within CpG islands of cancer cells relative to normal cells, or of the fact that in particular instances such methylation changes result in altered gene expression (or chromatin structure or stability), is inadequate to allow for effective diagnostic, prognostic and therapeutic application of this knowledge. This is

¹ Calculated as: [number of CpG sites / (number of C bases X number of G bases)] X band length for each fragment.

² Calculated as: (number of C bases + number of G bases) / band length for each fragment.

because only a limited number of CpG islands have been characterized, and thus there is insufficient knowledge, as to which particular CpG islands, among many, are actually involved in, or show significant correlation with cancer or the etiology thereof. Moreover, complex methylation patterns, involving a plurality of methylation-altered DNA sequences, including those that may have the sequence composition to qualify as CpG islands, may exist in particular cancers.

Therefore there is a need in the art to identify and characterize specific methylation altered DNA sequences, and to correlate them with cancer to allow for their diagnostic, prognostic and therapeutic application.

Summary of the Invention

The present invention provides for a diagnostic or prognostic assay for cancer, comprising: obtaining a tissue sample from a test tissue; performing a methylation assay on DNA derived from the tissue sample, wherein the methylation assay determines the methylation state of a CpG dinucleotide within a DNA sequence of the DNA, and wherein the DNA sequence is a sequence selected from the group consisting of sequences of SEQ ID NOS:1-103, sequences having a nucleotide sequence at least 90% identical to sequences of SEQ ID NOS:1-103, CpG island sequences associated with sequences of SEQ ID NOS:1-103, CpG island sequences associated with sequences having a nucleotide sequence at least 90% identical to sequences of SEQ ID NOS:1-103, and combinations thereof, wherein the CpG island sequence associated with the sequence of the particular SEQ ID NO is that contiguous sequence of genomic DNA that encompasses at least one nucleotide of the particular SEQ ID NO sequence, and satisfies the criteria of having both a frequency of CpG dinucleotides corresponding to an Observed/Expected Ratio >0.6, and a GC Content >0.5; and determining a diagnosis or prognosis based, at least in part, upon the methylation state of the CpG dinucleotide within the DNA sequence. Preferably, the DNA sequence is a sequence selected from the group consisting of CpG island sequences associated with sequences of SEQ ID NOS:1-103, CpG island sequences associated with sequences having a nucleotide sequence at least 90% identical to sequences of SEQ ID NOS:1-103, and combinations thereof. Preferably, the DNA sequence is a sequence selected from the group consisting of CpG island sequences associated with sequences of SEQ ID NOS: 2, 4, 6, 7, 9-16, 19, 20, 22-33, 35-43, 48, 51-55, 59, 60, 64, 71, 76, 78-81, 84 and 87-90, and combinations thereof. Preferably, the methylation assay procedure is selected from the group consisting of MethyLight, MS-SnuPE (methylation-sensitive single nucleotide primer extension), MSP (methylation-specific PCR), MCA (methylated CpG island amplification), COBRA (combined bisulfite restriction analysis), and combinations thereof. Preferably, the methylation state of the CpG dinucleotide within the DNA sequence is that of hypermethylation, hypomethylation or normal methylation. Preferably, the cancer is selected

from the group consisting of bladder cancer, prostate cancer, colon cancer, lung cancer, renal cancer, leukemia, breast cancer, uterine cancer, astrocytoma, glioblastoma, and neuroblastoma. Preferably, the cancer is bladder cancer, or prostate cancer.

- 5 CpG-containing nucleic acid comprising a carrier means containing one or more containers comprising: a container containing a probe or primer which hybridizes to any region of a sequence selected from the group consisting of SEQ ID NOS:1-103, and sequences having a nucleotide sequence at least 90% identical to sequences of SEQ ID NOS:1-103; and additional standard methylation assay reagents required to affect detection of methylated
- 10 CpG-containing nucleic acid based on the probe or primer. Preferably, the additional standard methylation assay reagents are standard reagents for performing a methylation assay from the group consisting of MethyLight, MS-SNuPE, MSP, MCA, COBRA, and combinations thereof. Preferably, the probe or primer comprises at least about 12 to 15 nucleotides of a sequence selected from the group consisting of SEQ ID NOS:1-103, and
- 15 sequences having a nucleotide sequence at least 90% identical to sequences of SEQ ID NOS:1-103.

- The present invention further provides an isolated nucleic acid molecule comprising a methylated or unmethylated polynucleotide sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:10, SEQ ID NO:12, SEQ ID
- 20 NO:13, SEQ ID NO:18, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:32, SEQ ID NO:34, SEQ ID NO:37, SEQ ID NO:38, SEQ ID NO:42, SEQ ID NO:44, SEQ ID NO:51, SEQ ID NO:52, SEQ ID NO:62, SEQ ID NO:64, SEQ ID NO:65, SEQ ID NO:68, SEQ ID NO:69, SEQ ID NO:70, SEQ ID NO:71, SEQ ID NO:74, SEQ ID NO:76, SEQ ID NO:82, SEQ ID NO:83, SEQ ID NO:84, SEQ ID NO:86, SEQ ID NO:90, SEQ ID NO:92, SEQ ID NO:97, and SEQ ID NO:100. Preferably the nucleic acid is methylated. Preferably, the nucleic acid
- 25 is unmethylated.

Detailed Description of the Invention

30 Definitions:

“GC Content” refers, within a particular DNA sequence, to the [(number of C bases + number of G bases) / band length for each fragment].

- “Observed/Expected Ratio” (“O/E Ratio”) refers to the frequency of CpG dinucleotides within a particular DNA sequence, and corresponds to the [number of CpG
- 35 sites / (number of C bases X number of G bases)] X band length for each fragment.

“CpG Island” refers to a contiguous region of genomic DNA that satisfies the criteria of (1) having a frequency of CpG dinucleotides corresponding to an “Observed/Expected Ratio” >0.6), and (2) having a “GC Content” >0.5. CpG islands are

typically, but not always, between about 0.2 to about 1 kb in length. A CpG island sequence associated with a particular SEQ ID NO sequence of the present invention is that contiguous sequence of genomic DNA that encompasses at least one nucleotide of the particular SEQ ID NO sequence, and satisfies the criteria of having both a frequency of CpG dinucleotides corresponding to an Observed/Expected Ratio >0.6), and a GC Content >0.5.

“**Methylation state**” refers to the presence or absence of 5-methylcytosine (“5-mCyt”) at one or a plurality of CpG dinucleotides within a DNA sequence.

“**Hypermethylation**” refers to the methylation state corresponding to an *increased* presence of 5-mCyt at one or a plurality of CpG dinucleotides within a DNA sequence of a test DNA sample, relative to the amount of 5-mCyt found at corresponding CpG dinucleotides within a normal control DNA sample.

“**Hypomethylation**” refers to the methylation state corresponding to a *decreased* presence of 5-mCyt at one or a plurality of CpG dinucleotides within a DNA sequence of a test DNA sample, relative to the amount of 5-mCyt found at corresponding CpG dinucleotides within a normal control DNA sample.

“**Methylation assay**” refers to any assay for determining the methylation state of a CpG dinucleotide within a sequence of DNA.

“**MS.AP-PCR**” (Methylation-Sensitive Arbitrarily-Primed Polymerase Chain Reaction) refers to the art-recognized technology that allows for a global scan of the genome using CG-rich primers to focus on the regions most likely to contain CpG dinucleotides, and described by Gonzalgo et al., *Cancer Research* 57:594-599, 1997.

“**MethyLight**” refers to the art-recognized fluorescence-based real-time PCR technique described by Eads et al., *Cancer Res.* 59:2302-2306, 1999.

“**Ms-SNuPE**” (Methylation-sensitive Single Nucleotide Primer Extension) refers to the art-recognized assay described by Gonzalgo & Jones, *Nucleic Acids Res.* 25:2529-2531, 1997.

“**MSP**” (Methylation-specific PCR) refers to the art-recognized methylation assay described by Herman et al. *Proc. Natl. Acad. Sci. USA* 93:9821-9826, 1996, and by US Patent No. 5,786,146.

“**COBRA**” (Combined Bisulfite Restriction Analysis) refers to the art-recognized methylation assay described by Xiong & Laird, *Nucleic Acids Res.* 25:2532-2534, 1997.

“**MCA**” (Methylated CpG Island Amplification) refers to the methylation assay described by Toyota et al., *Cancer Res.* 59:2307-12, 1999, and in WO 00/26401A1.

Overview

The present invention provides for 103 DNA sequences (*i.e.*, “marker sequences”) having distinct methylation patterns in cancer, as compared to normal tissue. These methylation-altered DNA sequence embodiments correspond to 103 DNA fragments isolated

from bladder and prostate cancer patients, and in many instances, represent novel sequences not found in the GenBank database. None of the instant sequence embodiments have previously been characterized with respect to their methylation pattern in human cancers including, but not limited to, those of the bladder and prostate. The significance of such methylation patterns lies in the value of altered fragments as potential prognostic, diagnostic and therapeutic markers in the treatment of human cancers.

Identification of Methylation-altered Marker Sequences in Genomic DNA

The MS.AP-PCR technique was used to scan the genomes of bladder or prostate cancer patients for DNA methylation changes relative to normal individuals, because the pattern is known to be highly conserved. A total of 103 DNA sequence embodiments (methylation-altered DNA sequences; "marker sequences") were isolated and characterized as having distinct methylation patterns in cancer, as compared to normal tissue.

Methods for the Identification of Marker Sequences in Genomic DNA. There are a variety of art-recognized genome scanning methods that have been used to identify altered methylation sites in cancer cells. For example, one method involves restriction landmark genomic scanning (Kawai et al., *Mol. Cell. Biol.* 14:7421-7427, 1994), another involves MCA (methylated CpG island amplification; Toyota et al., *Cancer Res.* 59:2307-12, 1999), and yet another involves MS.AP-PCR (Methylation-Sensitive Arbitrarily-Primed Polymerase Chain Reaction; Gonzalzo et al., *Cancer Res.* 57:594-599, 1997), which allows for a global scan of the genome using CG-rich primers to focus on the regions most likely to contain CpG dinucleotides. The MS.AP-PCR technique used in the present invention is a rapid and efficient method to screen ("scan") for altered methylation patterns in genomic DNA and to isolate specific sequences associated with these changes.

Briefly, genomic DNA from the tissue of bladder or prostate cancer patients was prepared using standard, art-recognized methods. Restriction enzymes (e.g., HpaII) with different sensitivities to cytosine methylation in their recognition sites were used to digest these genomic DNAs prior to arbitrarily primed PCR amplification with GC-rich primers. Fragments that showed differential methylation (e.g., *hypermethylation* or *hypomethylation*, based on the methylation sensitivity of the restriction enzyme, or upon DNA sequence analysis or Ms-SNuPE analysis; Gonzalzo & Jones, *Nucleic Acids Res* 25:2529-2531, 1997) were cloned and sequenced after resolving the PCR products on high-resolution polyacrylamide gels. The cloned fragments were used as probes for Southern blot analysis to confirm differential methylation of these regions in the tissue. Methods for DNA cloning, sequencing, PCR, high-resolution polyacrylamide gel resolution and Southern blot analysis are well known by those of ordinary skill in the relevant art.

Results. A total of 500 DNA fragments that underwent either hypermethylation (an increase in the level of methylation relative to normal) or hypomethylation (a decrease in the

level of methylation relative to normal) were isolated from the scanned patients genomic DNA. A total of 178 of these fragments were sequenced, of which 103 were *novel* in that they corresponded to DNA loci whose methylation pattern had not previously been characterized. The corresponding sequences are disclosed as [SEQ ID NOS:1-103], wherein
5 for certain sequences, the letter “n” refers to an undetermined nucleotide base.

Novel marker sequences identified by MS.AP-PCR. Table I shows an *overall* summary of methylation patterns and sequence data corresponding to the 103 DNA fragments identified by MS.AP-PCR. A total of 103 fragments were sequenced following identification as becoming either hypermethylated (gain of methylation; noted as having a
10 hypermethylation pattern) or hypomethylated (loss of methylation; noted as having a hypomethylation pattern) relative to normal tissue. For the fragments of each category, the “Average GC Content” is shown, calculated as (number of C bases + number of G bases)/band length for each fragment, as well as the average Observed/Expected Ratio (“O/E Ratio”), calculated as [number of CpG sites/(number of C bases X number of G bases)] X
15 band length for each fragment. Additionally, the percent of fragments that qualify as CpG islands is listed, and corresponds to the percentage of all fragments within each category that have sequence compositions that satisfy the criteria of having a “GC Content” >0.5 and an “O/E Ratio” >0.6.

Thus, of these 103 fragments identified by MS.AP-PCR, 60 showed hypermethylation
20 (Table I, upper row; Table II, [SEQ ID NOS:1-60]) while 43 showed hypomethylation (Table I, lower row; Table II, [SEQ ID NOS:61-103]). Moreover, 55 (43 hypermethylated, and 12 hypomethylated) of the 103 fragments correspond to CpG islands (*i.e.*, fulfill the criteria of a GC content >0.5 and an Observed/Expected Ratio >0.6;), whereas the other 48 (17 hypermethylated and 31 hypomethylated) fragments do not meet the criteria for CpG islands
25 (*see* Table II).

TABLE I. Summary of 103 DNA Fragments Identified by MS.AP-PCR

DNA Fragment Type	Methylation Pattern (relative to normal)	Number of Fragments (103 total)	Average GC Content	Average O/E Ratio	Percent that correspond to CpG Islands
Hypermethylated Fragments	Hyper-methylation	60	0.54	0.72	72%
Hypomethylated Fragments	Hypo-methylation	43	0.52	0.48	28%

Table II shows a summary of methylation pattern and sequence data for each
30 *individual* sequence embodiment ([SEQ ID NOS:1-103]), corresponding to the 103 DNA fragments identified by MS.AP-PCR. Data for the 103 fragments was divided into either hypermethylated ([SEQ ID NOS:1-60]) or hypomethylated ([SEQ ID NOS:61-103]) categories. Table II also lists, for each sequence embodiment, the corresponding “Fragment

Name,” fragment “Size” (in base pairs; “bp”), “GC Content,” Observed/Expected Ratio (“O/E Ratio”), “Description” (*i.e.*, as a CpG island if criteria are met), “Inventor Initials” (IDCM = Isabel D.C. Markl, JC = Jonathan Cheng, GL = Gangning Liang, HF = Hualin Fu, YT = Yoshitaka Tomigahara), “Cancer Source,” and “Chromosome Match” to the GenBank database. A dash (“-”) indicates that no GenBank chromosome match existed, or that only a low-scoring partial match was found. Averages of the “GC Content” and “O/E Ratio,” along with the percent of fragments that are CpG islands, are listed after the last member of both the hypermethylated and hypomethylated categories.

Therefore, the present invention provides for 103 DNA fragments and corresponding marker sequence embodiments (*i.e.*, methylation-altered DNA sequences) that are useful in cancer prognostic, diagnostic and therapeutic applications.

Additionally, at least 55 of these 103 sequences correspond to CpG islands (based on GC Content and O/E ratio); namely [SEQ ID NOS:2, 4, 6, 7, 9-16, 19, 20, 22-33, 35-43, 48, 51-55, 59, 60, 64, 71, 76, 78-81, 84 and 87-90]. Thus, based on the fact that the methylation state of a portion of a given CpG island is generally representative of the island as a whole, the present invention further encompassed the novel use of the 55 CpG islands associated with [SEQ ID NOS:2, 4, 6, 7, 9-16, 19, 20, 22-33, 35-43, 48, 51-55, 59, 60, 64, 71, 76, 78-81, 84 and 87-90] in cancer prognostic, diagnostic and therapeutic applications, where a CpG island sequence associated with the sequence of a particular SEQ ID NO is that contiguous sequence of genomic DNA that encompasses at least one nucleotide of the particular SEQ ID NO sequence, and satisfies the criteria of having both a frequency of CpG dinucleotides corresponding to an Observed/Expected Ratio >0.6, and a GC Content >0.5.

TABLE II. Summary of MS.AP-PCR Fragments Sequenced

Methylation Pattern	Fragment Name	Size (bp)	GC Content	O/E Ratio	Description	Inventor Initials	Cancer Source	Chromosome Matches	[SEQ ID NO]
Hypermethylation category	11-1A	510	0.44	0.74		IDCM	Bladder	-	1
	14-3B	313	0.58	0.74	CpG Island	IDCM	Bladder	2	2
	18-2B	165	0.57	0.45		IDCM	Bladder	7	3
	24-1B	601	0.51	0.72	CpG Island	IDCM	Bladder	Xp11	4
	26-1B	801	0.48	0.56		IDCM	Bladder	-	5
	26-2C	204	0.50	0.63	CpG Island	IDCM	Bladder	-	6
	30-3D	205	0.55	1.25	CpG Island	IDCM	Bladder	14	7
	32-3E	597	0.57	0.10		IDCM	Bladder	20q12-13.1	8
	34-2B	500	0.62	0.66	CpG Island	IDCM	Bladder	20	9
	34-4B	343	0.70	0.81	CpG Island	IDCM	Bladder	-	10
	34-5D	291	0.62	0.96	CpG Island	IDCM	Bladder	9	11
	34-6A	266	0.64	0.93	CpG Island	IDCM	Bladder	-	12
	35-1C	553	0.64	0.63	CpG Island	IDCM	Bladder	-	13
	36-2D	156	0.60	0.58	CpG Island	IDCM	Bladder	10	14
	38-1A	300	0.70	0.80	CpG Island	IDCM	Bladder	10	15
	38-2B	196	0.56	0.89	CpG Island	IDCM	Bladder	15	16
	7-8E	299	0.59	0.39		IDCM	Bladder	17q21-22	17
	83-4B	363	0.54	0.49		IDCM	Bladder	-	18

Methylation Pattern	Fragment Name	Size (bp)	GC Content	O/E Ratio	Description	Inventor Initials	Cancer Source	Chromosome Matches	[SEQ ID NO]
	84-1D	322	0.55	0.90	CpG Island	IDCM	Bladder	7	19
	101-3E	255	0.57	0.83	CpG Island	IDCM	Bladder	17	20
	M1-5A	406	0.45	0.96		IDCM	Bladder	1	21
	U2-8E	210	0.56	0.61	CpG Island	IDCM	Bladder	2	22
	U12-1A	310	0.56	0.81	CpG Island	IDCM	Bladder	2	23
	U7-4A	305	0.59	0.80	CpG Island	IDCM	Bladder	-	24
	NU9-5A	379	0.67	0.83	CpG Island	JC	Bladder	-	25
	3-17-8-B	625	0.48	0.72	CpG Island	GL	Bladder	18	26
	4-10-4-A	499	0.55	0.30	CpG Island	GL	Bladder	7	27
	1-1-1-A	561	0.58	0.98	CpG Island	GL	Bladder	20	28
	3-17-8-A	717	0.50	0.68	CpG Island	GL	Bladder	17	29
	G145-H	280	0.50	1.10	CpG Island	GL	Bladder	11	30
	1-1-1-D	270	0.50	0.60	CpG Island	GL	Bladder	2	31
	1-1-1-C	347	0.65	1.25	CpG Island	GL	Bladder	-	32
	G178-A	342	0.55	0.85	CpG Island	GL	Bladder	2	33
	34-A	370	0.62	0.44		HF	Prostate	-	34
	34-D	213	0.53	0.74	CpG Island	HF	Prostate	2	35
	35-D	173	0.56	0.66	CpG Island	HF	Prostate	3	36
	36-A	369	0.67	0.70	CpG Island	HF	Prostate	-	37
	40-A	123	0.60	1.16	CpG Island	HF	Prostate	-	38
	91-1	450	0.64	0.86	CpG Island	YT	Bladder	5 or 16q24.3	39
	93-2	593	0.51	0.68	CpG Island	YT	Bladder	Xp11	40
	93-3	457	0.52	0.94	CpG Island	YT	Bladder	Xp22.1-22.3	41
	94-8	211	0.66	0.96	CpG Island	YT	Bladder	-	42
	95-5	141	0.63	0.79	CpG Island	YT	Bladder	14	43
	97-5	559	0.56	0.40		YT	Bladder	-	44
	98-1	433	0.46	0.96		YT	Bladder	1	45
	100-1	487	0.59	0.58		YT	Bladder	14	46
	100-2	403	0.60	0.47		YT	Bladder	3	47
	100-6	155	0.57	0.99	CpG Island	YT	Bladder	20	48
	4-2	256	0.57	0.40		YT	Bladder	7	49
	5-8	224	0.47	0.96		YT	Bladder	5	50
	6-4	313	0.70	0.82	CpG Island	YT	Bladder	-	51
	7-6	385	0.70	0.88	CpG Island	YT	Bladder	-	52
	13-3	307	0.59	0.89	CpG Island	YT	Bladder	10	53
	15-2	182	0.62	0.92	CpG Island	YT	Bladder	13	54
	23-2	523	0.54	0.87	CpG Island	YT	Bladder	Xp22.1-22.3	55
	39-2	795	0.46	0.64		YT	Bladder	13	56
	40-2	438	0.62	0.51		YT	Bladder	10	57
	41-3	611	0.47	0.70		YT	Bladder	18	58
	105-4	291	0.58	0.71	CpG Island	YT	Bladder	5	59
	107-8	226	0.53	0.96	CpG Island	YT	Bladder	11	60
AVERAGE			0.54	0.72	72% islands				
Hypo-methylation category	14-2B	580	0.55	0.51		IDCM	Bladder	2	61
	16-1B	633	0.56	0.39		IDCM	Bladder	-	62
	18-1B	703	0.45	0.35		IDCM	Bladder	17	63
	19-1B	420	0.66	0.87	CpG Island	IDCM	Bladder	-	64
	20-1B	496	0.61	0.59		IDCM	Bladder	-	65
	21-2C	637	0.60	0.33		IDCM	Bladder	9q34	66
	29-1A	595	0.55	0.27		IDCM	Bladder	Xp11.23	67
	29-2B	580	0.47	0.77		IDCM	Bladder	-	68
	32-1A	589	0.59	0.48		IDCM	Bladder	-	69
	34-1B	450	0.42	0.46		IDCM	Bladder	-	70
	34-3B	432	0.70	0.61	CpG Island	IDCM	Bladder	-	71

Methylation Pattern	Fragment Name	Size (bp)	GC Content	O/E Ratio	Description	Inventor Initials	Cancer Source	Chromosome Matches	[SEQ ID NO]
	32-2B	748	0.47	0.24		IDCM	Bladder	2	72
	32-4B	599	0.57	0.15		IDCM	Bladder	20q12-13.1	73
	32-5B	614	0.58	0.20		IDCM	Bladder	-	74
	33-1A	552	0.54	0.32		IDCM	Bladder	10	75
	5-1E	501	0.61	1.04	CpG Island	IDCM	Bladder	-	76
	6-1A	826	0.55	0.36		IDCM	Bladder	22q13.32-13.33	77
	7-5D	433	0.59	0.85	CpG Island	IDCM	Bladder	5	78
	8-7C	424	0.58	0.83	CpG Island	IDCM	Bladder	5	79
	30-6D	285	0.63	0.72	CpG Island	IDCM	Bladder	1	80
	66-2E	401	0.54	0.82	CpG Island	IDCM	Bladder	16	81
	78-1C	268	0.54	0.41		IDCM	Bladder	-	82
	97-2E	989	0.53	0.16		IDCM	Bladder	-	83
	M1-8C	250	0.64	0.99	CpG Island	IDCM	Bladder	-	84
	M2-5A	402	0.50	0.45		IDCM	Bladder	5	85
	M1-4P	595	0.43	0.41		IDCM	Bladder	-	86
	M12-10A	304	0.53	0.76	CpG Island	IDCM	Bladder	7	87
	M12-12C	296	0.51	0.64	CpG Island	IDCM	Bladder	17	88
	M2-8M	220	0.67	0.62	CpG Island	IDCM	Bladder	6q27	89
	NU4-3A	273	0.63	1.02	CpG Island	JC	Bladder	-	90
	NU5-2A	361	0.44	0.73		JC	Bladder	6q14.3-15	91
	88-5	462	0.62	0.39		YT	Bladder	-	92
	90-1	591	0.66	0.45		YT	Bladder	19	93
	91-3	279	0.58	0.45		YT	Bladder	5 or 16q24.3	94
	91-4	351	0.55	0.30		YT	Bladder	18q23	95
	91-7	171	0.61	0.59		YT	Bladder	11	96
	89-3	743	0.55	0.43		YT	Bladder	-	97
	94-2	589	0.53	0.41		YT	Bladder	22q13.31-13.32	98
	94-3	538	0.53	0.49		YT	Bladder	5 or 18	99
	94-4	486	0.61	0.57		YT	Bladder	-	100
	94-5	450	0.60	0.45		YT	Bladder	1p36.2-36.3	101
	94-6	292	0.58	0.32		YT	Bladder	8 or 9	102
	96-4	395	0.63	0.54		YT	Bladder	9	103
AVERAGE			0.52	0.48	28% islands				

Diagnostic and Prognostic Assays for Cancer. The present invention provides for diagnostic and prognostic cancer assays based on determination of the methylation state of one or more of the disclosed 103 methylation-altered DNA sequence embodiments.

Typically, such assays involve obtaining a tissue sample from a test tissue, performing a methylation assay on DNA derived from the tissue sample, and making a diagnosis or prognosis based thereon.

The methylation assay is used to determine the methylation state of one or a plurality of CpG dinucleotide within a DNA sequence of the DNA sample. According to the present invention, possible methylation states include *hypermethylation* and *hypomethylation*, relative to a normal state (*i.e.*, non-cancerous control state). Hypermethylation and hypomethylation refer to the methylation states corresponding to an *increased* or *decreased*, respectively,

presence 5-methylcytosine ("5-mCyt") at one or a plurality of CpG dinucleotides within a DNA sequence of the test sample, relative to the amount of 5-mCyt found at corresponding CpG dinucleotides within a normal control DNA sample.

A diagnosis or prognosis is based, at least in part, upon the determined methylation state of the sample DNA sequence compared to control data obtained from normal, non-cancerous tissue.

Methylation Assay Procedures. Various methylation assay procedures are known in the art, and can be used in conjunction with the present invention. These assays allow for determination of the methylation state of one or a plurality of CpG dinucleotides (*e.g.*, CpG islands) within a DNA sequence. Such assays involve, among other techniques, DNA sequencing of bisulfite-treated DNA, PCR (for sequence-specific amplification), Southern blot analysis, use of methylation-sensitive restriction enzymes, etc.

For example, genomic sequencing has been simplified for analysis of DNA methylation patterns and 5-methylcytosine distribution by using bisulfite treatment (Frommer et al., *Proc. Natl. Acad. Sci. USA* 89:1827-1831, 1992). Additionally, restriction enzyme digestion of PCR products amplified from bisulfite-converted DNA is used, *e.g.*, the method described by Sadri & Hornsby (*Nucl. Acids Res.* 24:5058-5059, 1996), or COBRA (Combined Bisulfite Restriction Analysis) (Xiong & Laird, *Nucleic Acids Res.* 25:2532-2534, 1997).

COBRA. COBRA analysis is a quantitative methylation assay useful for determining DNA methylation levels at specific gene loci in small amounts of genomic DNA (Xiong & Laird, *Nucleic Acids Res.* 25:2532-2534, 1997). Briefly, restriction enzyme digestion is used to reveal methylation-dependent sequence differences in PCR products of sodium bisulfite-treated DNA. Methylation-dependent sequence differences are first introduced into the genomic DNA by standard bisulfite treatment according to the procedure described by Frommer et al. (*Proc. Natl. Acad. Sci. USA* 89:1827-1831, 1992). PCR amplification of the bisulfite converted DNA is then performed using primers specific for the interested CpG islands, followed by restriction endonuclease digestion, gel electrophoresis, and detection using specific, labeled hybridization probes. Methylation levels in the original DNA sample are represented by the relative amounts of digested and undigested PCR product in a linearly quantitative fashion across a wide spectrum of DNA methylation levels. In addition, this technique can be reliably applied to DNA obtained from microdissected paraffin-embedded tissue samples. Typical reagents (*e.g.*, as might be found in a typical COBRA-based kit) for COBRA analysis may include, but are not limited to: PCR primers for specific gene (or methylation-altered DNA sequence or CpG island); restriction enzyme and appropriate buffer; gene-hybridization oligo; control hybridization oligo; kinase labeling kit for oligo probe; and radioactive nucleotides. Additionally, bisulfite conversion reagents may include: DNA denaturation buffer; sulfonation buffer; DNA recovery reagents or kit (*e.g.*,

precipitation, ultrafiltration, affinity column); desulfonation buffer; and DNA recovery components.

Preferably, assays such as “MethyLight” (a fluorescence-based real-time PCR technique) (Eads et al., *Cancer Res.* 59:2302-2306, 1999), Ms-SNuPE (Methylation-sensitive Single Nucleotide Primer Extension) reactions (Gonzalzo & Jones, *Nucleic Acids Res.* 25:2529-2531, 1997), methylation-specific PCR (“MSP”; Herman et al., *Proc. Natl. Acad. Sci. USA* 93:9821-9826, 1996; US Patent No. 5,786,146), and methylated CpG island amplification (“MCA”; Toyota et al., *Cancer Res.* 59:2307-12, 1999) are used alone or in combination with other of these methods.

MethyLight. The MethyLight assay is a high-throughput quantitative methylation assay that utilizes fluorescence-based real-time PCR (TaqMan®) technology that requires no further manipulations after the PCR step (Eads et al., *Cancer Res.* 59:2302-2306, 1999). Briefly, the MethyLight process begins with a mixed sample of genomic DNA that is converted, in a sodium bisulfite reaction, to a mixed pool of methylation-dependent sequence differences according to standard procedures (the bisulfite process converts unmethylated cytosine residues to uracil). Fluorescence-based PCR is then performed either in an “unbiased” (with primers that do not overlap known CpG methylation sites) PCR reaction, or in a “biased” (with PCR primers that overlap known CpG dinucleotides) reaction. Sequence discrimination can occur either at the level of the amplification process or at the level of the fluorescence detection process, or both.

The MethyLight may assay be used as a quantitative test for methylation patterns in the genomic DNA sample, wherein sequence discrimination occurs at the level of probe hybridization. In this quantitative version, the PCR reaction provides for unbiased amplification in the presence of a fluorescent probe that overlaps a particular putative methylation site. An unbiased control for the amount of input DNA is provided by a reaction in which neither the primers, nor the probe overlies any CpG dinucleotides. Alternatively, a qualitative test for genomic methylation is achieved by probing of the biased PCR pool with either control oligonucleotides that do not “cover” known methylation sites (a fluorescence-based version of the “MSP” technique), or with oligonucleotides covering potential methylation sites.

The MethyLight process can be used with a “TaqMan®” probe in the amplification process. For example, double-stranded genomic DNA is treated with sodium bisulfite and subjected to one of two sets of PCR reactions using TaqMan® probes; e.g., with either biased primers and TaqMan® probe, or unbiased primers and TaqMan® probe. The TaqMan® probe is dual-labeled with fluorescent “reporter” and “quencher” molecules, and is designed to be specific for a relatively high GC content region so that it melts out at about 10 °C higher temperature in the PCR cycle than the forward or reverse primers. This allows the TaqMan® probe to remain fully hybridized during the PCR annealing/extension step. As the

Taq polymerase enzymatically synthesizes a new strand during PCR, it will eventually reach the annealed TaqMan® probe. The Taq polymerase 5' to 3' endonuclease activity will then displace the TaqMan® probe by digesting it to release the fluorescent reporter molecule for quantitative detection of its now unquenched signal using a real-time fluorescent detection system.

Typical reagents (e.g., as might be found in a typical MethyLight-based kit) for MethyLight analysis may include, but are not limited to: PCR primers for specific gene (or methylation-altered DNA sequence or CpG island); TaqMan® probes; optimized PCR buffers and deoxynucleotides; and Taq polymerase.

Ms-SNuPE. The Ms-SNuPE technique is a quantitative method for assessing methylation differences at specific CpG sites based on bisulfite treatment of DNA, followed by single-nucleotide primer extension (Gonzalzo & Jones, *Nucleic Acids Res.* 25:2529-2531, 1997). Briefly, genomic DNA is reacted with sodium bisulfite to convert unmethylated cytosine to uracil while leaving 5-methylcytosine unchanged. Amplification of the desired target sequence is then performed using PCR primers specific for bisulfite-converted DNA, and the resulting product is isolated and used as a template for methylation analysis at the CpG site(s) of interest. Small amounts of DNA can be analyzed (e.g., microdissected pathology sections), and it avoids utilization of restriction enzymes for determining the methylation status at CpG sites. Typical reagents (e.g., as might be found in a typical Ms-SNuPE-based kit) for Ms-SNuPE analysis may include, but are not limited to: PCR primers for specific gene (or methylation-altered DNA sequence or CpG island); optimized PCR buffers and deoxynucleotides; gel extraction kit; positive control primers; Ms-SNuPE primers for specific gene; reaction buffer (for the Ms-SNuPE reaction); and radioactive nucleotides. Additionally, bisulfite conversion reagents may include: DNA denaturation buffer; sulfonation buffer; DNA recovery reagents or kit (e.g., precipitation, ultrafiltration, affinity column); desulfonation buffer; and DNA recovery components.

MSP. MSP (methylation-specific PCR) allows for assessing the methylation status of virtually any group of CpG sites within a CpG island, independent of the use of methylation-sensitive restriction enzymes (Herman et al. *Proc. Natl. Acad. Sci. USA* 93:9821-9826, 1996; US Patent No. 5,786,146). Briefly, DNA is modified by sodium bisulfite converting all unmethylated, but not methylated cytosines to uracil, and subsequently amplified with primers specific for methylated versus unmethylated DNA. MSP requires only small quantities of DNA, is sensitive to 0.1% methylated alleles of a given CpG island locus, and can be performed on DNA extracted from paraffin-embedded samples. Typical reagents (e.g., as might be found in a typical MSP-based kit) for MSP analysis may include, but are not limited to: methylated and unmethylated PCR primers for specific gene (or methylation-altered DNA sequence or CpG island), optimized PCR buffers and deoxynucleotides, and specific probes.

MCA. The MCA technique is a method that can be used to screen for altered methylation patterns in genomic DNA, and to isolate specific sequences associated with these changes (Toyota et al., *Cancer Res.* 59:2307-12, 1999). Briefly, restriction enzymes with different sensitivities to cytosine methylation in their recognition sites are used to digest genomic DNAs from primary tumors, cell lines, and normal tissues prior to arbitrarily primed PCR amplification. Fragments that show differential methylation are cloned and sequenced after resolving the PCR products on high-resolution polyacrylamide gels. The cloned fragments are then used as probes for Southern analysis to confirm differential methylation of these regions. Typical reagents (*e.g.*, as might be found in a typical MCA -based kit) for MCA analysis may include, but are not limited to: PCR primers for arbitrary priming Genomic DNA; PCR buffers and nucleotides, restriction enzymes and appropriate buffers; gene-hybridization oligos or probes; control hybridization oligos or probes.

Kits for Detection of Methylated CpG-containing Nucleic Acid. The reagents required to perform one or more art-recognized methylation assays (including those identified above) are combined with primers or probes comprising the sequences of SEQ ID NOS:1-103, or portions thereof, to determine the methylation state of CpG-containing nucleic acids. For example, the MethyLight, Ms-SNuPE, MCA, COBRA, and MSP methylation assays could be used alone or in combination, along with primers or probes comprising the sequences of SEQ ID NOS:1-103, or portions thereof, to determine the methylation state of a CpG dinucleotide within a genomic sequence corresponding to SEQ ID NOS:1-103, or to CpG island sequences associated with sequences of SEQ ID NOS:1-103, where the CpG island sequence associated with the sequence of the particular SEQ ID NO is that contiguous sequence of genomic DNA that encompasses at least one nucleotide of the particular SEQ ID NO sequence, and satisfies the criteria of having both a frequency of CpG dinucleotides corresponding to an Observed/Expected Ratio >0.6, and a GC Content >0.5.

SEQUENCE LISTING

<110> Markl, Isabel
Tomigahara, Yoshitaka
Liang, Gangning
Fu, Hualin
Jones, Peter

<120> Methylation Altered DNA Sequences as Markers Associated with Human Cancer

<130> 47465-14

<160> 103

<170> PatentIn version 3.0

<210> 1

<211> 510

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> ()..()

<223> "n" refers to an undetermined base

<400> 1

ttgcaagccc ccttngctct tcctttgncc tcgcctacat attcagggga tcgcaatctc 60

actcgcgaaa taatttnttt ctgtaagagg aagccgcctt tcccctctcc caccgccaag 120

gtaaaggctg ctaaagtagc tcttcttgga aggaaaaata ttttaaaaag cagctggggtt 180

gctctccaca agaagatggc agttttggga aaaccatta tgtgtccaaa tgccggtttc 240

cttttcttgt ttaacgcttt ttagagggca aaaatgacgc tcatgtgaag cccacaggct 300

cgagccaatg tcgctgggct aattatgagt ctgcttatcc cactcccaaa tatccgagac 360

gactcactca naagacattt ttactcttcc aagaattgng aattcagaan cagcttcccc 420

acattctaag agaaaaaaaa acttgtttaa cgggcacggt tttgattttt ttgccgctgg 480

cgaccttaat taaaagccgg gagctncnna 510

<210> 2

<211> 313

<212> DNA

<213> Homo sapiens

<400> 2

gcactcttaa aacgcctctc tgcagtccca ggtccgcgct cccaagaac tggccagatc 60

gcgcggggct tggccctga caactctgcc tctccacct gttgcgttta ctccgtttag 120

ttggctgtgc agtctctggc cccaggtgtg cttttaaaac tcgaggaacg cgggtgttg 180

actcattcgc agcctcttgc ctctggttcc cgtgatccca cgggtggcgag cttccaggct 240

cagcgaggag atctgggttt gaacattcat ctcccatggt actcttttct tgcctctcgc 300
 gtccccaagc cga 313
 5
 <210> 3
 <211> 165
 <212> DNA
 10 <213> Homo sapiens
 <400> 3
 gcttagcaaa tttccctttt ttattgttgg ttttgtctgt tggctcttac cccctttcct 60
 15 tttcctgctt cccctgagtc agcaatgctg agcccagcga agcacagggg gccaaaggga 120
 gagacacacg gagcgccccg gggccccca gcctcggcgg ccaaa 165
 20 <210> 4
 <211> 601
 <212> DNA
 <213> Homo sapiens
 25 <400> 4
 ggggggagtc gtgcgtgtca gatttaggcc aggaagcggg agtcgccagc agcgagagtt 60
 taacctctgt gggcgagag ggttgcgggg attcagcgcc cgggaccgtg gatctgtgca 120
 30 gggagtcata ggtgtgtgtg acatcagtgg tggaacattt tggctcgttt tcacaattca 180
 gtcattatcc tttctgcttt cctcctggaa gcattaaggt tgaagttttc ttctaaagat 240
 caaagttttg atttgttata ttagttcgga tttgtttgat ttttgtttgt gttcggtttc 300
 35 aagtgtgat ttgtaacttt tctccccccc cacacacacg ccttttgacc cctgaattat 360
 ttaaaagtcc attgttgag tggaacaacat cctccgagac tcaaagggca aggccatggg 420
 40 cgctttattc cggctgctgc tccaggaacg tgggaaagca gcggagtttt attctagggg 480
 aaggaaacaa aggcggccga gtgccagctg cacgtttggt gggatttggt catcaggggt 540
 ggacatgctg cccaatggag ctgtcggcag tttgaccag cttggtccgt cgcgtccga 600
 45 a 601
 <210> 5
 50 <211> 801
 <212> DNA
 <213> Homo sapiens
 <220>
 55 <221> misc_feature
 <222> ()..()
 <223> "n" refers to an undetermined base
 60 <400> 5
 gagtacgcgg ggcagaacca gcgcaataca gcattctggt aggggaacta attttgacta 60

5
 10
 15
 20
 25
 30
 35
 40
 45
 50
 55
 60

aaatatttgc caattctaata ccccaattcc tggacctccg ggtagctggc aaggtatttt 120
 atgttagatg tgtctggagt aaggtgcacc ggagtatttc gacaagagac tcaattcaat 180
 gcgtattaaa acttgattga gagagggaga gagagaggtc attttataaa gaaagacctg 240
 tgaacactgt agattggaaa tttatgtttg caaaataaaa ggatgggttt atcaagtgga 300
 tgcatttaca aaatgtggca tccaggtttc gtaaaattag ctgaattcta cgggtaagat 360
 tatgaatgtg gctcataaat aattaatagg tagtgaaaaa gaatgtattt tgcattaggc 420
 agtgcattca atagtatttc ggaaatgagc acttcgattt cctcggnttc catgcgnggc 480
 cacctctcca gagcagggca ggcacccagg gnggtgcccc cacaacaag cgcgtgtggg 540
 cattttcttg gctcgtgcgc tgaagtgcac gctgggcctt ggtgcccgc caetcagcct 600
 gggagatagg gaggtgggtc tacctgcagg ccgatttgtt ccccgccata ggacactagt 660
 gggcggcaaa cctcacaaga ctcttcagc cagccttcag cagagccagc aaaccagcc 720
 gccaccgagg gaggactgct ccatgcagat ggtcaggggc tttcttctga agacgcctcc 780
 cccacgatct ctcaagttca c 801
 <210> 6
 <211> 204
 <212> DNA
 <213> Homo sapiens
 <400> 6
 ccggttgctg ctggaggatg ggactacgaa ggatggggac tccgctcggc caccgctcct 60
 gaatggcctc taatctcggg gttaaatact ttatgagagt atcaatacca cctaatectt 120
 tgetgagaat tactgctaga aatgtagatt ctgaggttcc gaaagtttgt ttttggttac 180
 cccctccagc tctccccgcg gcaa 204
 <210> 7
 <211> 205
 <212> DNA
 <213> Homo sapiens
 <400> 7
 agacctttat cgggcgtgag tgaaatcggg cgttcgtgtt tttcgtgggt cttccaacca 60
 caggccgcct gagatggttc tagtcccttt gagtatacag acccttcctg tgcattgacc 120
 gacacagctc ggcccggatc ccgaaatgaa cgtttctacc ttcggaacgc tgcgtctcgg 180
 atccttctga acccgcacgt cgcaa 205
 <210> 8
 <211> 597
 <212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

5 <222> ()..()

<223> "n" refers to an undetermined base

<400> 8

10 gaccatgaaa tcgtgtggct ctagccccc ttctgggcctct tgttggtaat gaagccactc 60

taaagcgccc cctgttattc agagggctcc ccagctgcc a gatatgtgt atggggaggg 120

15 catagcaggt ccttttgccc cggcagccat tcttctgctc acaaggggct ggctctgggg 180

acagggatgt ctttgtcatc agtgaccact aatccccctc ctcatgggcc tccagggtctg 240

ctccccctca ctctcttggt tgaagttgta ggggctgagg ttaccctgag aaacacctgt 300

20 tcttgagacc catagacca accttgaga tgcaggggga gccactggct gggtctgca 360

ngtggggcca gctgatcccc anctgctggc acctccaggc atccacagag cttggagtcc 420

25 cagccacatt tcctccttgg ccttagaggg agaggaagtc ctttgattgc ctagtccaag 480

atccctttat ttctgcctt gggattatgg ggnagcaagc catgcccttc atgggaagct 540

gttctccctt cctcgggggt gggctctggc tcagctcggg caacagtcac gatgggc 597

30

<210> 9

<211> 500

<212> DNA

<213> Homo sapiens

35

<220>

<221> misc_feature

<222> ()..()

<223> "n" refers to an undetermined base

40

<400> 9

gccaaacgcn ataccctctg cggggtgaga atgcggggcc gcccggtcc tcccgtgagg 60

45 ccagggcctc ctgttctcct agacaccca aggagccaac tcctccgcag aagttccccg 120

cttctgctct tatttccaag cttegcgctt tctacaaact ccctgttgcc ttgactttga 180

50 tttccagccg tggtaggggt cagagtgaac cccggcgcgc tccccgacgg catccccgca 240

caccaggata ggagaaattg gagggcctgg ggcctcgggc tccgcagtcg tcggaggaag 300

aaccaccgc ggggtccgca agggaaagtg aagaggcccc ggatttttcc aaagcgctgg 360

55 ccaggacccc gaaggaaggg gaggagtcac ctgaagccgg ggaaggcccc ttgggtgctc 420

tgccttggat ccttatgttc actgactttc gcgaccctg gaggggggca aatccgcgct 480

gtttccccc acttggttc 500

60

<210> 10
 <211> 343
 <212> DNA
 <213> Homo sapiens

5

<400> 10
 gccaaaccac accagtacct gggaccgggg ggagcccgt cggccgcta aaccgggctg 60
 gctggcgcca gggctccggg aggtgcggtc cggcggggaa gccgtgatgg gaagcgactc 120
 10 tgtccaggga gtgtccttca ccaccacact cctcacgtcc aggcagtgat cgacggcctg 180
 gcggcaccct cacagcgggc ccatagcacg gggccacaca cgtcccctga gcttagcctg 240
 15 ggcacattcg tctgccgccc agggcttaag ccagtctgca gcccgcgcc cgtcactcgg 300
 acgcaagtcc gtcgtccgct ctgccacgcg gccgctaagc cga 343

20

<210> 11
 <211> 291
 <212> DNA
 <213> Homo sapiens

25

<400> 11
 gtcttacaca ctccgcacac aacgcggccg gtgttaagtc tccaaacgcc ccgagagctc 60
 caaggaccgc gcgcgcgaag gcgccgtagc aagtgggcac acaccagaca ccaccccggc 120
 30 gtgttccgcy ggagaagcca gtgcacacat cctcccgcaa ggcgggggtg ccagtgcac 180
 acaggaatcc tgcccttttt ctagaaaagc cccctcccc actttccctc caatacactc 240
 acctgcgtct caacagtttc cttcttgccg tacacgcggc cgctaagccg a 291

35

<210> 12
 <211> 266
 <212> DNA
 <213> Homo sapiens

40

<400> 12
 gtccggatca gtttccccgg ccaggteget tcccggtctc aaccatttcg cgctctgctc 60
 45 tgtccgctgg tttgtccctg cccggttcct ctccccgggc ctgtcagcct ccgcttctct 120
 ggaggttcct gggactcatc tctgatccac cgtcttgctg tctctgggcy catcgacttc 180
 tctccatctt cgggctcact cctgactccc tcgctgccgc ccccgggggt ttccacgcgt 240
 50 gtctctaacc gcggccgcta agccga 266

55

<210> 13
 <211> 553
 <212> DNA
 <213> Homo sapiens

60

<220>
 <221> misc_feature
 <222> () .. ()

<223> "n" refers to an undetermined base

<400> 13

5 gatcctggtc catcgaaacc ttgtgtgcat cggtagtgcc ttcttgggcg tttgtttcta 60
gcccagcctg acagtggagt gccagaaaga gggagaggac cgtcatggct actctgcccc 120
10 tgggtgtcacc atgcgctctc ccccggcacc ggcgaggcga aacgtttcgc tagtcccccg 180
gaggccccctc ggtcagggca gcagcatccc tgcaccctct ccgcagggtg tctccccgac 240
gccacaggtg gccagcaggc cgcgggtggg ggcaggagcg cctctcccct gcccaggcct 300
15 cccgctcctt ctcggagcgc tgtggcgggg tggagagaca gccttctaca gctagtctag 360
ctcggcgcgg tccccgtctg tggcctccta atcccacagc cacagcgctt tcctctaacc 420
tccctcgggtg ggcttaaagc ctcccgttcc ttctgtctca ttctttctgc tccctcccc 480
20 cgaaaccccc agatganagc tgggaacctg gcncantna ctgagcnaac agtggtgacg 540
ggccgnggcc caa 553

<210> 14

<211> 156

<212> DNA

<213> Homo sapiens

<400> 14

30 gcgcacacag tgggtacaag gatgagctcg gtgtaaggaa tggaaagccc ccagtctaaa 60
ccaccgcccc ctagacacgg gtgaaaacct gcctaaaagc taactcaggc agtgactcta 120
35 tcacccgaag gggccctggg ccgcggccca agccga 156

<210> 15

<211> 300

<212> DNA

<213> Homo sapiens

<220>

45 <221> misc_feature

<222> ()..()

<223> "n" refers to an undetermined base

<400> 15

50 gttcacagcc cataaggtgg ggggtggccc aacctgaaac ggagcctgag ccaggatcct 60
gcaaccaaag tctgaagcgc ccccgggtgg gggccgagag cgctgcaggc aggtggnggc 120
55 gcggggcagg cgggcgggcg aaggagctc cggntacga ganaacgcgg agcgccccct 180
tcccacctgc gcgagggcat cctgcccggg ggaggaaagg cgggagtcg aggcgggtcg 240
60 gattcccagc cagctccctc ctcacaggag gcggccatt atccggcgtc gcaaagccga 300

<210> 16
 <211> 196
 <212> DNA
 <213> Homo sapiens

5

<400> 16
 ggcgccagc aggggagcga gggaggagg tgcagaaaga ggctccgaaa ttgggggaaa 60
 ctgacccgtg cttctctacc ttcggagggtg ggacagttgc acgaagtgt agttagaccg 120
 gatcagttgg aactgacgga ggactgcaaa gaagaaacta aaatagacgt cgaaagcctg 180
 tcctcggcgt cgcaaa 196

10

<210> 17
 <211> 299
 <212> DNA
 <213> Homo sapiens

15

<220>
 <221> misc_feature
 <222> ()..()
 <223> "n" refers to an undetermined base

20

<400> 17
 acaccaggag aggggaagaa nccagcacct accgacaggg gtggagctgg gtcaagaatg 60
 gtgtgggtccc tgctttgggg gaatgtctgg gaggtagaaa gcccttcta acggggcgctc 120
 actgcaatta ctgcttctc tttcccataa aactccccct agtgtatcag aaccccccaag 180
 gagtttcagt aagcggttct tctgttgtct ccggtgaga ctccagggga acctcaagct 240
 cacatggccc tggccgggcc cctgggcagg agcaggcgag aggtctgcgc ggccgctaa 299

25

<210> 18
 <211> 363
 <212> DNA
 <213> Homo sapiens

30

<400> 18
 gggatatgtg tacacatccg agataactac acaggcatcg accctgtcca cccggggatg 60
 ctagaggggc tgcgtggtt ttactccagg ccatgggtgag agccaccgtg aacacagggc 120
 tctctctct gagctgcaga agctctgtgc cctgtccct gccacaagtc acagactttc 180
 ttcatgtgtt ttacctcatg ttaatgaagg agatcttctc caggggcttg atctagtggg 240
 aaacagagga gggggggatt ttaaatttca gtccgtccaa ccctgtagat ctgctgtcct 300
 acagtaacgt aaaggatcac caggtaaaac gctgcttctc ccggacgccg ccccgaagc 360
 cga 363

35

<210> 19
 <211> 322

40

<210> 19
 <211> 322

45

<210> 19
 <211> 322

50

<210> 19
 <211> 322

55

<210> 19
 <211> 322

60

<212> DNA
<213> Homo sapiens

<400> 19

5 ccggcccgtc cctcttaata tggcctcagt tccgaaaacc acagaataga accgcggtcc 60
tattccatta ttcctagctg aggtatccag gcggctcgga cctgctttga acactctaata 120
10 tttttcaaag taaacgcttc gggctgcagg acactcagct aagagcatca ggggggcgcc 180
aagaggcaag gggcggggat ggggtggtggc tcgcctcgtg gcagaccgcc cgcccgtcc 240
caagatccaa ctacgagctt tttaactgca gcaactttaa tatacgctat tggagctgga 300
15 attaccgcgg ccgctaagcc ga 322

<210> 20
<211> 255
<212> DNA
<213> Homo sapiens

<400> 20

25 taataagata ccaaatacggg cgagaaacga aaagctcctg gcctccgtat ttggggccag 60
agacaccgca gggagtcagg tccccgccga caaatcggaa gaggcctgcg ggagttagcc 120
agataatgct ctccctgtcc taccgcgtccc caccaatttg ccttttacct gccgcagagc 180
30 ttgcttgaac caaagggggt tgcgggtcttc tcctcctcaa cttgcgatcc ccaggccttc 240
gcgtcccgaa gccga 255

35 <210> 21
<211> 406
<212> DNA
<213> Homo sapiens

40 <220>
<221> misc_feature
<222> ()..()
<223> "n" refers to an undetermined base

<400> 21

45 atgtgnaag gctcgctntc catttctctt ttctccttc tccctctctc atgtgcggtc 60
50 tccctcaaca tccaaaccaa ccgagtgcgt ctgaggtgaa atcgtgccag acttagagac 120
ggctgccagg tttctctcaa gtcttggtt aacaaaagaa agcaaattac aaaaatggaa 180
attttcaaac tagcgttcag tggattcaa atcgacgttt gggtagcgca caggcacaga 240
55 ccgcattcgt gctattttgt gattaaaatg ataccaaaaa tacctccttg ctttggtttt 300
cgtcttcgaa aacgacttct ttctctcttc taatttcccc cttacttttg ggagcggcaa 360
60 acccctgacc actctagaat tgctaacatt tggaccggcg tcgcaa 406

5 <210> 22
 <211> 210
 <212> DNA
 <213> Homo sapiens
 <220>
 <221> misc_feature
 <222> ()..()
 <223> "n" refers to an undetermined base
 10
 <400> 22
 gcacgttcgn gcnnctgtga ccatnagctg ccaactggan gcaccnnggn aaggggtgggg 60
 15 gcctcctgga gacttngggg agagggatag ccgntaaag ctctgtcct ttctataggc 120
 ataagcgggt ggtcaccacg gattggggat cccgaatccc tggctccaga tagacttaat 180
 20 gaagaagcac ctggatccgg gccgcgncaa 210
 <210> 23
 <211> 310
 <212> DNA
 25 <213> Homo sapiens
 <220>
 <221> misc_feature
 <222> ()..()
 30 <223> "n" refers to an undetermined base
 <400> 23
 35 tcacgcttnc naaggctctg aatcctgagg gncagatctc caagaaggag ggaggctggt 60
 cctagttooc gaggtcctnn actaggtcta gatcactggg taaaagaagg ggagcggcan 120
 cacgtatggg gtaggcgctc tcactactca catctcgaga cctttgccgg cgtagggctg 180
 40 tccgggggga acgaccgccc ttttcggta tcggttgta tggcggcgcc cagcccagcc 240
 tgggtttttc cggtagccaa ttgaactaac aaccccgctc cctttaggac taatctgtca 300
 45 cgtcggcgca 310
 <210> 24
 <211> 304
 <212> DNA
 50 <213> Homo sapiens
 <220>
 <221> misc_feature
 <222> ()..()
 55 <223> "n" refers to an undetermined base
 <400> 24
 60 ctctgggtctg tgntggatac gcgtgttctt ctgcggagtt aaagggtcgg ggacgggggt 60
 tctggactta ccanagcaat tccagccggt gggcgtttgg cagtcactta aggaggtagg 120

gaaagcagcg agcttcaccg ggcgggctac gatgagtagc atgacgggca gcagcagcag 180
 5 ccagcaaaag ccctcgcaaa gtgtccagct gctgcactgc cgcggggact cccacagcac 240
 catgactagt tcgtgcgact ctgcancanc aaacggcttc cgaggaacac angatcgcg 300
 gggca 304
 10
 <210> 25
 <211> 379
 <212> DNA
 <213> Homo sapiens
 15
 <220>
 <221> misc_feature
 <222> ()..()
 <223> "n" refers to an undetermined base
 20
 <400> 25
 aaaacncatn tgnagagcnc ntcggcagag ncgagctgg ctgacccagg agaaggcgcg 60
 25 ctgggtgtgg ctgggacggc caaggccgcg gcttcccgcg tggggatgcg ctntggcgca 120
 aagctgggtcc cggcggggcc aggcgtttgt gggcgggtga cggggatcta gggcttccgc 180
 30 tcnggattcc tcttgggctg tctttncggg tttggactcg cctgccaggc tgtgtgcagg 240
 gttcccgcgtg cctctggccg gcaggcgctc gggctgcagg tgggccggca ggcaggtggt 300
 agcgggaagg gagcacaggt agcgaggtgg gatcggcgac ctggctaggg tgtcggcaga 360
 35 atggaatgcg cggccgcta 379
 <210> 26
 <211> 625
 40 <212> DNA
 <213> Homo sapiens
 <220>
 <221> misc_feature
 45 <222> ()..()
 <223> "n" refers to an undetermined base
 <400> 26
 50 gggacgcna g cagggantt tgatccgttt tgaatgaaaa gaaagagaan ccaaaccaaa 60
 cctntcagtc atccaaaacc ttcaggcttc caggaggtt ttgctataat tttctctaag 120
 55 catgactggt tctgggggag gggaaagggg tggttgtatt tactgaaaat tcaaatcgaa 180
 ataataaatg gccaaatttg gacacttacg gacccaaaca gttttgctca cgccagagaa 240
 accgagagca cagggttgc gtgaagccta tctcggcaga aggcaacatt ctaataaagc 300
 60 ccgtgggaaa acagattaca ttttcgccat gaataagtca tgcagtgaaa aatattgcct 360

acagcctgtc gacttatatt attatcacgt ttttcaactc ggcgtgagga gggagaggag 420
 tgttcatatt tgactaggaa ttgcaggatc gatgcaaact ccagggcagc agccagactg 480
 5 gcatatgtgg ggctctccgg ttacttttctc tgtatgtcgc gggtgagagg aacagcgagg 540
 acaatttagc gcaaacacac gaagggtcgg atctcaaggg ggcagcgctg ggagaaaggt 600
 tagggctgna gagcgnanag ncaa 625
 10

<210> 27
 <211> 499
 <212> DNA
 15 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> ()..()
 20 <223> "n" refers to an undetermined base

<400> 27
 25 gnctccnctg tcccctcggg cggaacggag gcaactttcc ggagtctatt tttgttaaga 60
 caatcaactc caataactga gctgaagttt ttgttttaaaa agaaaaaat ctgataagtg 120
 atgattttac ctacttgtgg aactagatt tcaattagga aggttttttt aaacggcttt 180
 30 ttgtaacttc gctgcaggaa gcaggtttgt ttctttttct tttcttttta agagaagggtg 240
 tatttcaactg gtgcaatggc ttggcacctc cggggcctgg gaggacctca gacctcccca 300
 gccctggggt tctccgtctt caagaccaac taggaagggt caagcgggga gagggagtgg 360
 35 agggtcagggt gagatctcag agctgccccg gccggcccc gtctctttct acctctctt 420
 ccagagaacc agcggctcac acccttctca acgcaggaca tgctcggcgg ccaaagccga 480
 40 attctgcaga tatccatca 499

<210> 28
 <211> 561
 45 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 50 <222> ()..()
 <223> "n" refers to an undetermined base

<400> 28
 55 gggcgattgt tattcaaacn ngntanctct ctgcggggnn gagnaatgng ggcctcgcac 60
 ggctncatcc ccgtcgagen cagggcctcc ctgttctnct agacatncca aggagccaac 120
 60 tctccgcag aagttccccg cttctgctct tatttccaag cttegcgctt tctacaaact 180
 ccctgttgcc ttgactttga tttccagccg tgggtagggt cagagtgaac cccggcgcgc 240

tccccgacgg catccccgca caccaggata ggagaaattg gagggcctgg gcctcggctc 300
 5 cgcagtcgt cggaggaaga acccaccgcg gggccccaa gggaaagtga agaggcccg 360
 gatTTTTCCA aagcgtgcc aggaccccgga aggaagggga ggagtcacct gaagccgggg 420
 aagctccttg ggtgctctcc ttggatcctt atgttcaactg actttcgoga ngccccctgg 480
 10 agnggaaaaa tccgcgtgt tcccccaac ttaacttcac gcggccgcta agccgaattc 540
 tgcngaaatc attacactng c 561

15 <210> 29
 <211> 717
 <212> DNA
 <213> Homo sapiens

20 <220>
 <221> misc_feature
 <222> ()..()
 <223> "n" refers to an undetermined base

25 <400> 29
 actctccgcg gtntcntggt gcctcacagg aggtggggct cctccaccc ggtccccagg 60
 cctctccctc tgcccagact tcccggtcct gcctccttcg cctcgctgc ctgcccagact 120
 30 ctgaaccctg ctctcttct aactaaaagt cagtgtttta tttcctccgc agtccaatgc 180
 ccgcgtttta cttattcaa taagaagggc ttcatTTATg gcaagacagg acagccaggt 240
 35 aataagggcc tctgcacacg cgggccatt ggaggggagg aactgcgaag tcttcccgga 300
 agagcttctt ggagagaagg ggaacgagcc agcgtttatt gagcatctat tatactaagc 360
 atctgcttgg cagttcacga cggtcgcatt ttttcatcct tacagcgatc cctattgtgt 420
 40 cgcttgcttt aaagcctcac agtcacaaa gggtgggat ttattccaga tctctctctc 480
 agatgccatc tcacttccag gtgtctctgc tgctttgaac gcgggaaacc cacgcaaagg 540
 45 agtgatttcc aaggccttct gtttgaata tctttaatcc tccccttatt aactggaaaa 600
 actcccacgc atccttcagg gctcagctca aatgtccttt atntctgcag ngaaactttc 660
 ccaaggaaaa ttagttacac agctaatttt agataaattg agccagttga tagaatt 717

50 <210> 30
 <211> 280
 <212> DNA
 55 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> ()..()
 60 <223> "n" refers to an undetermined base

<400> 30
 tgatggatat ctgcagaatt cgggctttgn gacgccgggc acgcagtagg gaaaacagta 60
 5 ttaaaacgcc ctacagaaaa tctcggcgaa gtccccggag aactctgggt tctaagatca 120
 gctggggcgca ctttctccgg gacgtccctt cttctcggtc tcagcgccct cctgccctca 180
 10 gccgcgcng tnttgttttg gtggcaaact gaaataagaa atggaaatat attggccttt 240
 gctgctgccca gggatgagag gttgttgacg tcggcgcaaa 280

15 <210> 31
 <211> 270
 <212> DNA
 <213> Homo sapiens

20 <220>
 <221> misc_feature
 <222> ()..()
 <223> "n" refers to an undetermined base

25 <400> 31
 gnggnngna nncggcgatg gatntnngna gantnngtg atggatatct gcagaattcg 60
 gcttagcggc cgcaacaaa gagcgaacca aaggatgctt cgaattttta aaacggaatc 120
 30 tctgcacca aatgcaggac tgggtactta aggagctgcg aagtctgatt taccgggcct 180
 actctcgacc tgccccccac ccccgactca gggggacctt tttatcntga acgccagagc 240
 35 tacnnaccaa gtcgggtggc cacnnccaaa 270

40 <210> 32
 <211> 347
 <212> DNA
 <213> Homo sapiens

45 <220>
 <221> misc_feature
 <222> ()..()
 <223> "n" refers to an undetermined base

50 <400> 32
 tttggannta ngggggcgtg gcgtggatcc agtttcccc ggccaggten gcttcccggt 60
 ctcaaccatt tcgcgtctg ctctgtccgc tggtttgtcc ctgcccggtt cctctccccg 120
 ggctgtcag cctccgcttc tctggaggtt cctgggactc atctctgac caccgtcttg 180
 55 cgttctctgg gcgcacgac ttctctccat cttcgggctc actcctgact ccctcgctgc 240
 cgccccgggg gtttccacgc gtgtctctaa ccgcggccgc taagccgaat tctgcagata 300
 60 tccatcacng aantctgcag anatncatcg nccaannnca ccgcact 347

5 <210> 33
 <211> 342
 <212> DNA
 <213> Homo sapiens
 <220>
 <221> misc_feature
 <222> ()..()
 <223> "n" refers to an undetermined base
 10
 <400> 33
 gtagggcgcc gccgtgacag attagtccta aaggggaacgg ggttgtagt tcaattggct 60
 15 accggaaaaa accaggctgg gctggggcgcc cgccatgaca accgataccg gaaaaggcgg 120
 gtcgttcccc ccggacagcc ctacgccggc aaaggtctcg agatgtgagt agtgagagcg 180
 20 cctaccccat acngtcggcc ggctcccctt cttttacca gtgatctaga cctagtctag 240
 gacctcggga actaggacca gcctccctcc ttcttgga tctgaccctc aggattcann 300
 nntttgtctc acgagctcca acccnacnca tccaaannnc aa 342
 25
 <210> 34
 <211> 370
 <212> DNA
 <213> Homo sapiens
 30 <220>
 <221> misc_feature
 <222> ()..()
 <223> "n" refers to an undetermined base
 35
 <400> 34
 cattgtttac tttcgtctaa acgcggtgga agcccatgga agaaagcggg tagcagcaag 60
 40 gcagagccct gtcctctctg cagccccagc tcccagcgcc ctgggctttc caggcacctg 120
 tccgggtagg ggattgaggg ccgtggccag gccgcactt tctgctagc cgcagctggc 180
 45 cacatgcca tctgaccctc cgagttctcc tctaaaaatg gggctgacag ccgctacctc 240
 acaaagtcca caccgggctc aaccgntgc cttcctcccc aacaggactc tgccaccctc 300
 cctcaggatg cctgagggcc ccganctgca cctggccagc cantttgtga atgaggcctg 360
 50 nggggcgntt 370
 55 <210> 35
 <211> 213
 <212> DNA
 <213> Homo sapiens
 <220>
 <221> misc_feature
 60 <222> ()..()
 <223> "n" refers to an undetermined base

<400> 35
 5 aaaatacnan taaagcgatg cttcgaattt ttaaaacgga atctctgcac ccaaagtcag 60
 gactggtgac ttaaggagct gcgaagtctg atttaccggc ctactctcga cctgcccccc 120
 acccccagct caggggacct tttgtctgaa cgccagagct actgaccagg tcgggggggcc 180
 10 gcggcccaag ccgaattctg cagatatcca tca 213

<210> 36
 <211> 173
 15 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 20 <222> ()..()
 <223> "n" refers to an undetermined base

<400> 36
 25 gacnncgggt ttgtgtgtaa cagggtcagt ccccgatatct actttgcgaa agcttcgagg 60
 cgagcgtgaa gtcaagggtc gcggtggatg ggggtaaaan gcctcctcnt cccactgcct 120
 30 gcnccgtctt ggggtaaccc ctanccccca cccgngttn cnctttaatg etc 173

<210> 37
 <211> 369
 35 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 40 <222> ()..()
 <223> "n" refers to an undetermined base

<400> 37
 45 tcactgtgcc gggctctctc tncccggtcc aactccotta cttgtctctca tctctgtccc 60
 caaggctcgt gacccgcgga ggtgatgggg gggataggag agccccaggg accgcagagg 120
 tgacacaatc gcccgcctgt cctccctcgc tgggagccga ttcagcctgt gccgagcctc 180
 50 tcccttcgcg tgctctcgcg cacagcgggtg gcaccgcagg actccgggtc cccccggct 240
 ctccatcggg aagccggcaa atgcgcttcc tcagccagac cgcggcgggg tgggggcggg 300
 55 gggggcgga gttgaaatac tgggacagaa acacctgcc gtcccaagg acggaact 360
 ggatgcaa 369

<210> 38
 60 <211> 123
 <212> DNA

<213> Homo sapiens
 <220>
 <221> misc_feature
 5 <222> ()..()
 <223> "n" refers to an undetermined base

<400> 38
 10 gtcccttcgc cccgcttttn ctttcccna ggtcccagcg nccgaaccgg cggatgtcca 60
 cgaaacatag ggcgagccgg gggccangcg gggcgtgta aaatctcntg tggtcatttt 120
 gtg 123
 15

<210> 39
 <211> 450
 <212> DNA
 20 <213> Homo sapiens
 <220>
 <221> misc_feature
 <222> ()..()
 25 <223> "n" refers to an undetermined base

<400> 39
 30 ctagccctgg aagagaatcc gaggtcagc cntgctgcag caccaggac actgcatccc 60
 agcacctgcc cgaagatcag ccaggaccc aaaggaaagc aggtccaag ctccccggaa 120
 gccaaaggaaa ataggaaaac atatcctgcc ccggggacac cttctggaac tatgaccaca 180
 35 tgcacttgac cttccggaac aatcaccgca tgcacctgac ctcccgaac tgtcaccacc 240
 gcgcgcacct gacctcccg gactgtcacg accgcgcgca cctgacctcc cggcactgtc 300
 atcaccgcgc gcacctcacc tcccggaaact gtcaccaccg gcgcgcacctg acctcccggc 360
 40 actgtcacga ccgcgcgcac ctgacctccc ggaactgtca tcaccaggcg cacctgaccc 420
 ccgggcactg tcacgaccgc gcgcacctca 450
 45

<210> 40
 <211> 593
 <212> DNA
 <213> Homo sapiens
 50 <400> 40
 ggaccaagct gggtaaaactg ccgacagctc cattgggcag catgtccacc cctgatgacc 60
 aaatcccacc aaacgtgcag ctggcactcg gccgcctttg tttccttccc ctagaataaa 120
 55 actccgctgc tttccacgt tcttgagca gcagccggaa taaagcgccc atggccttgc 180
 cctttgagtc tcggaggatg tttgccactc caacaatgga cttttaaata attcaggggt 240
 60 caaaaggcgt gtgtgtgggg ggggagaaaa gttacaaatc agcacttgaa accgaacaca 300

aacaaaaatc aaacaaatcc gaactaatat aacaaatcaa aactttgatc tttagaagaa 360
aacttcaacc ttaatgcttc caggaggaaa gcagaaagga taatgactga attgtgaaaa 420
5 cgagccaaaa tgttccacca ctgatgtcac acacacctat gactccctgc acagatccac 480
ggtcccgggc gctgaatccc cgcaaccctc tgcgcccaca gaggttaaac tctcgctgct 540
10 ggcgacttcc gcttcttggc ctaaattctga cacgcacgac tccccccgcg gca 593

<210> 41
<211> 457
<212> DNA
15 <213> Homo sapiens

<400> 41
accaccaacc aaatagggcc tttcctgtta acgaccacgc ggcaaggggg ccggggccctc 60
20 gcacgcctcg acggcctccc ccactccaaa gggactccga tttcgcagga tctcccgctt 120
cccgctcttg ctcccaacac cctacgtttt tctcttctc ctcatttacg tatttacaat 180
25 aaaacagcga agctgcacag tctgtctcta aatcaaacgc ggttaccatc aaagcctcag 240
actctatgtc tcaaccgcaa aaggtctgac aggaaatcaa ctcgggagtt tgtcaattct 300
ttaaactcaa agctctgtta acgaaatctg gatctttcct cgctccccac ctgcctcccc 360
30 tgacaggaga atgactgtaa aaggatcctg tcgtccccga aagtcagcac caagcacttc 420
acaaattgtc aaatctcaaa agcttacacg cgcggca 457

35 <210> 42
<211> 211
<212> DNA
<213> Homo sapiens

40 <400> 42
gcctgacctg aatgacgcgc atgttgaggc cgggtctctg cgccagctgc tcgcggatgt 60
ggcgggtggg cttgggtgta gcagcgaagg cggccttcag cgtctccagc tgcttggtt 120
45 tgatgggtggg gcgcgggtccc cgccgcttgg cgcccagggt ctgggtcgtca ttctcgttgc 180
taccgcttc cttgtccgac acgtcggcgc a 211

50 <210> 43
<211> 141
<212> DNA
<213> Homo sapiens

55 <400> 43
aaatcatctc cgggggcccc gcacggacac gctocagacc cgtgagttcc ccagcgcct 60
gccgggaggt caggggcgct gaaagaagga agaattcagc cacctctcag catccctgtt 120
60 acctcgagga cgcgcctctc a 141

<210> 44
 <211> 559
 <212> DNA
 5 <213> Homo sapiens

<400> 44
 acccactttc cattaacact aaataaaacg catccatgga tttcctctcc attccgaggc 60
 10 aacaggagtg catggcacat tgcctactc ccctgaagct cttcgctaac ctaagactcc 120
 agggtgagga agttagctgg agctttttta agtgcatctc caaagagaat tttgctcaca 180
 15 ccatgagagc cccaagaaa caccagggcc cccttagatg ccggagacca cgccctccag 240
 gaataagccg caccctctgc ccagcagatc cttgcgcgag tagccctctt tccctggggc 300
 taatcaagtg catgccacat gtcaccactc tcagctggca attcttctc agaggcgcag 360
 20 actttcacgg aatccccagc aggggggggt aagagattca ggggaggccc cgcccggtgcc 420
 ttcacaaaaa gtcgctttac cgtggctcgt gtctgcggc cccaaggggg tagcctggga 480
 25 cgtgtattgg gagggcatag aggctcctc caggacaagc tgccagcctc cagtgggcaa 540
 ccatgtgaga ggcaaaatt 559

<210> 45
 30 <211> 433
 <212> DNA
 <213> Homo sapiens

<400> 45
 35 gcgaacagca caaaggcttc attcctacga gagattaagt tttagagcaa atggacacga 60
 tcgttaaaga atttgatatt tccatgtaaa ctgcattagc aggttatgcg atccaaactc 120
 40 acaggaacaa ctccaactct cggccatgcc ctatttcatg tctagatttg ttttaaccgac 180
 ttacatcata atccaagaat acgaactaca gtatattctt acagcaaagt tattccttaa 240
 aagcaaaacc gagccacctt tgaaaacacg cacacacatt atccacggca ctaaaacccc 300
 45 agtcttgacc gagaaagacc aacaacttgg gggggaagaa aacaacttca gagccagagc 360
 tcccaaagca gaaagcgctg gcggctgaag ggcacacgag gttccgctcc cgggcgaacg 420
 50 ggcggcgctg caa 433

<210> 46
 <211> 487
 <212> DNA
 55 <213> Homo sapiens

<400> 46
 cccttagtat tccatgagcc accatthttcc ccacgatccc tccagcctga acgatcacat 60
 60 cctactgtgg accacgactc tcccagcagc gggcggtttta tatccagtta gcaggttctc 120

accacccccct cgctggctcg aatacagcat ctgcaccgag ttccccgagaa tcgtcaaccc 180

agcaaattccc ttaattgggtg gacatgaaaa tccagggcctt tgtgctgtaa taacagagtc 240

5 ctggggggcct ggggagtttg tgccgcttgg agctcaggtt tctgggacag aggctgagcg 300

cagggcaggg aggcaggtct cacctggcac ctcccagagt cctcgccgag cagatggaag 360

10 cagaggctct cgcgcccggc ccccgccggg agacctctct ctctttccct cggcctgctc 420

tgccctctcc cgccttctcc ctgtctgac cttctctgct gtcattgttct ttgtcctcgc 480

gccccga 487

15

<210> 47

<211> 403

<212> DNA

<213> Homo sapiens

20

<400> 47

gtcatataag cacaaccatt cccagggcca ccctggatgc atcagatcag tccccccact 60

25 ggtgaccaca atggctggct cagagtgcct ttgaacagac aggagaaaca gacttcttgg 120

agggagggac cttccacag ggaatggcca aggagctagg tcttcagggc ttgcatggcg 180

tggagtgtgt gctcaggtgc acagtgaagc aaacctgagg ggacttgggc cctgcgtcct 240

30 ccagcacaca cgcacccttt cgccgtcaca tccggggcac ccaccgctgg aatatgtgag 300

cgcacttgg ccagccacga gttccagggc caggaagtcg tgettctcgt tcaggcgccc 360

35 gttgtagaag agcagcccgc tctgctgcac tgtcgcgtcc cga 403

<210> 48

<211> 155

<212> DNA

<213> Homo sapiens

40

<400> 48

ggcgtggaga ggagggggca gaaactcagc cgcccctacg tttgctaaac tgcgtccgcc 60

45 agggggcgta tttttctaaa acgcacaaga cgtttcgtgg gttatcgatg gtctcttgag 120

cctccttgac tgatggggat tgaccgggag ggata 155

50

<210> 49

<211> 256

<212> DNA

<213> Homo sapiens

55

<400> 49

tctactgagc ttttctttaa gtggaaccag aagtgtggg atgagaggga aaggatggga 60

gtgcgtccaa aggtggacag caggtcccca tccctgggtg gagtgagact ggacggcatc 120

60 ccccggaag gtggtttggg ccttgacaaa ggctagaggc aggagtccat gatgcagaga 180

tgaacacagtg cccctccgcg tgtgagtcca cgaaggtcac tactgagget ttgtgottgt 240

aaaaggccgc cccgca 256

5

<210> 50

<211> 224

<212> DNA

<213> Homo sapiens

10

<400> 50

tgcggggctcg tgggggaacc ggcgggagct gttcgtggc cggcctcact ggagtaggaa 60

15

ttttagatga aactgagtcc gtttctcctt gaaggcaggc agtattctta gatctactat 120

tcatttaaaa agaaggaaaa gaaaaaaaaa tgactgctac ttactgagaa gaaaatttct 180

gttctcctcc gattccgctg atcccgcttt atccgcgcac ctca 224

20

<210> 51

<211> 313

<212> DNA

<213> Homo sapiens

25

<400> 51

gtggctggga cggcccaggc cgcggcttcc cgcgtgggga tgcgtgtgg cgcagagctg 60

30

gtcccggcgg ggcaggcgt ttgtgggcgg gtgacgggga tctagggctt ccgctcgtga 120

ttcctcttgg gctgtcttcc cgggtttgga ctgcctgcc cggctgtgtg cagggttccc 180

gctgcctctg gccggcaggc gtccgggctg cagggtggcc ggcaggcagg tgttagcggg 240

35

aaggagcac aggtagcgag gtgggatcgg cgacctggct aggggtgtcgg cagaatggaa 300

tgcgcggccg cta 313

40

<210> 52

<211> 385

<212> DNA

<213> Homo sapiens

45

<400> 52

tacgttgccg attcattctg ccgacaccct agccggtcgc cgatgccacc tcgctacctg 60

tgctcccttc ccgctaacac ctgcctgccg gccacactgc agcccggaag cctgccggcc 120

50

agaggcagcg ggaaccctgc acacagccgg gcaggcgagt ccaaaccggg aaagacagcc 180

caagaggaat cagcagcgga agccctagat ccccgctacc cggccacaaa cgctggccc 240

55

cgccgggacc agctctgcgc cacagcgcat cccacgcgg gaagccgcgg cctgggcccgt 300

cccagccaca ccagcgcgc cttctccagg gtcagccagc tgcggctctg ccgaagcgct 360

cctccgctcc tttctcgcgc cccga 385

60

<210> 53

<211> 307
 <212> DNA
 <213> Homo sapiens

5 <400> 53
 aacccggctc gggttcggcaa gggttcaggga gacaaggtag agaaggctgg ggtgagcaag 60
 aagtcggggcg gccgatcgtc agggccacga gcctcgctt gccttcttgg aatcccaccc 120
 10 aactttaaag gcccaaagat cctgaaaatt ccgaaagcga aactgcgggc tgggtctccag 180
 aagtttgaga acgggtctccc aggttttcca gcgtcgctccc gggattctcg gacaccacaa 240
 acgccatcaa ccacgagcac cgggtgtccgt ggctattgcc ccgaatggtc cccatccgcg 300
 15 tccccta 307

20 <210> 54
 <211> 182
 <212> DNA
 <213> Homo sapiens

25 <400> 54
 cgatgtcgaa gccgtttgga gggaacagcg gtttccaagt tcctgctgac ttgagaagcc 60
 tctgcggggtt tccgaatctc cggcgcactc ctgggcgcgc tgcgggagct gtagctcagc 120
 cagccaggga gtagcggctt tcatccgccg ggaggagtct ttcgagttca atcgcggggg 180
 30 ca 182

35 <210> 55
 <211> 523
 <212> DNA
 <213> Homo sapiens

40 <400> 55
 tcggggttga tccgccccaa ccaaataagg cctttcctgt taacgaccac gcggcaaggg 60
 ggccggggccc tcgcacgcct cgacggcctc cccactcca aagggaactcc gatttcgcag 120
 gatctcccgc ctcccgcctc tgctcccaac accctacgtt tttctcttcc tctcattta 180
 45 cgtatttaca ataaaacagc gaagctgcac agtctgtctc taaatcaaac gcggttacca 240
 tcaaagcctc agactctatg tctcaaccgc aaaaggctctg acaggaaatc aactcgggag 300
 50 tttgtcaatt ctttaaaactc aaagctctgt taacgaaatc tggatccttc ctgctcccc 360
 acctgcctcc cctgacagga gaatgactgt aaaaggatcc tgcgtcccc gaaagtcagc 420
 accaagcact tcacaaattg tcaaattctc aaagcttaca cgcgcgggca ctccgaaaag 480
 55 gctgtgggga ccacccaaag caccctcctc cacaccgcgg gca 523

60 <210> 56
 <211> 795
 <212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

5 <222> ()..()

<223> "n" refers to an undetermined base

<400> 56

10 tttacttttct tccggctgac gtccatctcc tcaaatttct caggaatgtg gggaagctcc 60
tagccctgcc tgcctttcta gagggcttct tggatttgca gcttctaaca agttgctctc 120
gccacggaga agctgttatt atgacaaaat atttggggca ttatcaaaat cacacaggct 180
15 gctgggctgc tgtcggtttc tcgccagggc cagtaagcag ttacatttgg agttgctacg 240
tggtgtttgg gggccgggct gtggagagtg actgagccag tatttttcat ccaaaattct 300
20 gcaaattgaa ttaaccacaa ttctagtctc acctcccgct tttaaaaaaa taagttgaag 360
aaaaggtaaa tattagagat aaggcagcat ctagtgactg cggagaggca caagctggtg 420
ggcgagggtt gggggagtca gcaaagccct tcaaaacctc cccgtttaat tttctggctg 480
25 tctctgcac cgtttgccag aattccaaat gcttgagtc atttanaggt gcgagaactc 540
aaacgtcggt ccacttgaa aggggaccgt ttaacgttaa attccattag cacctaaatt 600
30 gtttcttaaa gacatccgct cagacacagg actcgaaagc gagcatttca tgcaataaaa 660
tttctcaaat tttaaacctt gttaaaagct tgtctcgac ctcggtccc tcccctccc 720
cggaaganaa caataggccg ntggcgcatc cccacttcgg antaaatatt gacgggggaa 780
35 gttgctaaaa acatc 795

<210> 57

40 <211> 438

<212> DNA

<213> Homo sapiens

<400> 57

45 gcctgtgtgt aggggactgg aggtggggga acctgttctt ttcttgtgtc tgatcctggg 60
gctcgcttcc tgggtcctag aacagcagcc aggacggaag aaactgttca cgttgacccc 120
ctttctctaa gattcccagg ccaagagtag ctgcagaagg tggccctgaa tctatggcct 180
50 ccttctctct gcctgaccg gctagtggat ccggagaggg gaccaggag agtcctccg 240
agcaggggtc cttcgggaga cagagagggg tccaggctga gagaactctt caagcatggc 300
55 gagtctgcgt tatagaatcg ggcgggcggc tcaacttggg ggaagcacca agaagagctg 360
ggcgacctgg agcgcagaac cggctttggg gagccacccg gcggggcagg ggtagcacgg 420
agcccggggc gcggccca 438
60

<210> 58
 <211> 611
 <212> DNA
 <213> Homo sapiens

5

<400> 58
 gcttccccct tcctttctcc cgcgctgccc ccttgagatc cgacccttcg tgtgtttgcg 60
 ctaaattgtc ctcgctgttc ctctacccg cgacatacag agaaagtaac cggagagccc 120
 tacatatgcc agtctggctg ctgccctgga gtttgcacg atcctgcaat tcctagtcaa 180
 atatgaacac tcctctccct cctcacgccg agttgaaaaa cgtgataata atataagtcg 240
 acaggctgta ggcaatattt ttactgcat gacttattca tggcgaaaat gtaaactggt 300
 ttcccacggg ctttattaga atgttgccct ctgccgagat aggcttcacg caagccctgt 360
 gctctcagtt tctctggcgt gagcaaaact gtttgggtcc ataagtgtcc acatttggcc 420
 atttattatt tcgatttgaa ttttcagtaa atacaaccac ccctttcccc tccccagaa 480
 acagtcatgc ttagagaaaa ttatagcaaa acctccctgg aagcctgaag gttttggatg 540
 actgagaggt ttggtttggt ttctctttct tttcattcaa aacggatcaa actccctggc 600
 tcgcgtcccc a 611

30

<210> 59
 <211> 291
 <212> DNA
 <213> Homo sapiens

35

<400> 59
 gagtttggca ggccccggat tccacaaagg agtaggcgcg gccagccgcc tccagccctg 60
 agctcagtaa attcgggtgc ctgaatgctc ccttcctgtc cttaccactg cgagctctct 120
 tgggacagct ttctaggttc cactgcgacc tactttccgc tccctgagtg cttctttgct 180
 gaaactgcag gcgaaaagat ctctttccca gaccgcagcg cactttgaga aggggctcaa 240
 agtcgcccgc tctgaatccg gcaccggcaa ataggagtag ccgcatgcgc a 291

45

<210> 60
 <211> 226
 <212> DNA
 <213> Homo sapiens

50

<400> 60
 gaaaacagat aaaacgccct acagaaaatc tcggcgaagt cccggaggac tctggtttct 60
 aagatcagct gggcgcactt tctccgggac gtcccttctt ctcggtctca gcgccttcct 120
 gccctcagcc gcgcgcagct ttgttttggg gacaaaactga aataagaaat ggaaatatat 180
 tggcctttgc tgctgccagg gatgagaggt tggtgacgct ggcgca 226

60

<210> 61
 <211> 580
 <212> DNA
 <213> Homo sapiens

5

<400> 61
 ctgtgatgca ctccggcggat ctccggtggca gctgcctcct tcatctccag tgacgcctgc 60
 atgctgtcct aggcagtgtg aggagtgaag atgagatttg gcgcattcttt caacggagtc 120
 tgagcaaagc taaagggctc cgattcgtgc aagccaaggc ctgccccctcc tatcctgtcc 180
 tccttgagga cctgtgctaa ggctttctca tccaccaggc caccatgggc tgcgttcaca 240
 aggaatgctc cctgtctcat ctgctttata gtaaagtcac tgacgaggtg gtggttatgt 300
 tcattgagat tgctgtgcaa cgagacacag tcaactctgat acagcaaacc ctgcaggggtg 360
 tatcaggggc ccctctgcat gccctgggac ctctctatct tgtcctacaa gttaggggtca 420
 taaaatacga cgctgaatcc aaaggccttg gctcaaactg caaccgcctg cctcatgcaa 480
 ccgaagccca tgaggcctag cgtcttccac gaatgagggc cactcccatg gccacctcga 540
 gaatctgctc cagcgtctga acccgcgcac ctcaagccga 580

<210> 62
 <211> 633
 <212> DNA
 <213> Homo sapiens

30

<400> 62
 gcccaggaga agccctccac ggtgggcgtc ctccctagaca accagcacc cctgcaggca 60
 ccctcgtctg gcagaatcag ccctttccca cctgcaggcc cttctcagcg cctctgactt 120
 cccacacaca gcacaggta caaactggc cctggcagtg cactctagcg ggcctctctc 180
 acaagttctg cgggcctcgt ttcattggaaa gcgggttggtg gattcctgct gcccttgat 240
 ggccccctgcg cagcacacc tctgagcggg cactgagcga gcgtggggag ctgctccctg 300
 ggaactaggc aggagctttt aaacaccctt acacacagcc attctgcggg aatacatgct 360
 ttcccggtaa ggcttttact gttcattcca ggtaaattgg aagtcgcaca cccaagctc 420
 caaatacaac tcgttagctg gcaggctctt gaagccaatt ccttctgagg aaaatggaga 480
 taatagcagc taccctccca ggtgactggg ggagaataaaa gtggtctgtc atagtgggtg 540
 ttgcagctgg tggctgctat tatccttcat tacagcttgt aaaaagggtg tctaggccat 600
 ttacacacag ataggccggg tggggtaagc cga 633

55

<210> 63
 <211> 703
 <212> DNA
 <213> Homo sapiens

60

<400> 63
gcctatgaat ggatttataa ttgctttatt tttgtcccat ttagacagaa gtcagagaca 60
5 gaggagagaa ccaaaaaact tggatgtttc cgtaaactag attcgtcaat cctcgataat 120
tgaaagtagt tccagtatgt cagccaccgg gggtccctgg ggagctaacc agtcctgaag 180
gaagtatgaa gaggaagagg aggtcttcag ttaaggggat gaatttgtgc agtcctaagc 240
10 cctgcaaagg tgctggaggg aggaagaagg gcaggaaata aaagatggaa gaaaatttgt 300
tttttatcca cttagagttt tatctttaat gatgggaaac agtgctgctc tcaggaaact 360
15 cagtgtggag atctaggagt tcacggttca tagtccatta ggagcaggaa aaggatagag 420
gacatttata aagtaacatc caagtccaaa gtaaaatggg ataaattgtt tcccatgata 480
aaggctggct gagtaggtca ggaaaggctc tgtcagacca tatgtgctgt ttcaggctgc 540
20 ttcaaattct tttaggacag tgggtgatat gagtgaagac ggggcaggca ggccacatct 600
cttagaagag gaaggtgatt gccacgtctc cttcctccat gctgatggca aggcgtgcgg 660
getgtgttct cttgcagcca gcgtcccatg ctcggcggcc aaa 703

<210> 64
<211> 420
<212> DNA
30 <213> Homo sapiens

<400> 64
gtgacgtgcg gaatacacgt gatgtcgggg acaggagcgg gctgaagagg gcacgatccc 60
35 acgcgaggcg caccctcac ccggggtagg agcccgctgc acttgctgtc gctcagcccg 120
ggcgtgcac cacggcagcc gccatgctgc ttaaagccgg tcatgtgacg cgggagccag 180
40 ggtggaaggg gtccccgagg gcaagccttc gacacgtgac ctgccacccg actacggaag 240
cctcttgggc gttccgcccc ggctcacatg tcatgtgacg gccggccggg cgcgggagta 300
accaggaact ttccagacc ctgcggtccc tggagcgtca aaaagagcgt ccccgtagt 360
45 aggtggagtc gcctgccctt ccgaatctca gctgtcttat ctggaacccc cacgcgga 420

<210> 65
50 <211> 496
<212> DNA
<213> Homo sapiens

<400> 65
55 gcgctgcacc aatttagagg gtagaaaaag gaggtagaag caaagaggaa aaaataaata 60
aacaggcaac aaaaaccaa ccagccagc ctgagccatt tgcattagtg ttcatttagg 120
aaattagcag acgggaaacg ctggggagtg gaggggccc cggccttggg gactgcagag 180
60 cccgctcagc cctgggtggc tgggccaca tggctgtcgc caggagcaca ggaggacca 240

gaggtggccg agggagcctc gccgggctcc ggtatgggtc ctggcccctc acaggtgcca 300
gcctggccca gtgactgtgg acgctgtggg agagcaggcc tccgatacgc agggctggga 360
5 ctgctgacct ggaaggtggt gccgggctg tctggtgaag gcgccgttgg cagctagaga 420
gagacggcgg atggggtgac gccataacct acggtcccag ttttgaggct tgacggtgac 480
ggaaaaggac gtcggc 496

10

<210> 66
<211> 637
<212> DNA
15 <213> Homo sapiens

15

<220>
<221> misc_feature
<222> ()..()
20 <223> "n" refers to an undetermined base

20

<400> 66
cgccgagccg ggatgagcaa ggcttcctgg aggagagggc cggcctgagc ttggaaggat 60
25 ggggaggagc cactggctac aagggtgtag aggtgagaac cagtgtgacc tgcccatcgc 120
tggtcgtctc tgggtcattc agctgaaatg gcatctctga gctgagagga gtgttgccctg 180
30 taaggagcta ggcacagccc cccagtagag gggcgccca ggcacagccc atagccgcag 240
acttagtgag tctagctagg gagacagtag aggggcaaaa atgaggacac aggtcaccaa 300
aaatcctggc caggtcctgc cactacctgg ctacagcacc tgcccccccg agcctcagtt 360
35 tccccattg gtggaatgga gtgaggaaga cgcgcctccc ggggctgcca tggagaattg 420
agtcagagtc tgggggtgct gggagggctg gggagcagcc tccctgagcc tcagtttccc 480
40 tggttgggga atgaggacct tgctcgtccc ccctcataag ggaagctgt caggaaagtg 540
ctttcaacgc tgagccattt cccagtggtg cacaattagc tttccagagg attttggtgg 600
attctagagc tngagggctg ggggatnggc ggccaaa 637

45

<210> 67
<211> 595
<212> DNA
50 <213> Homo sapiens

50

<400> 67
gccctgagct cttgagggcc tctgcagttc ttgggacaat tctgggacta tatctttggg 60
55 ccttggtgag atctagaggc tctaaagtct ttgggagggg tcctgagctc cgtggacggc 120
agggctcttg gcactcactt gcattcttga ggggtgtgtt tggcctcgtc cgtgcaggctg 180
tagaatttcc cctgtagaga ggatgtctgt caagtaggtt cacccttcat cacactcccg 240
60 cccagacccc tgctggcat tccctccagt gtttgcccca ccttgaagag ctgcaccccg 300

60

atgcaggcga acataaattg cagaagtgtg gtgacaatca tgatgtttcc gatgggtccgg 360
 atggccacaa atacacactg caccacatgc tgcgggcacc caagcatatg gctactgaac 420
 5 actacaggcc acagtgggtca tggggcaggg actctgtgtca tagatgcagc tgagggactt 480
 gggctgggga catgtggtga tgggtcaggg atgtatggtt agcaacatgt gttcaagagg 540
 10 cagtgttatg ggctagagac gtgtgggcat ccaccaggaa taagtgtttg ccggg 595

<210> 68
 <211> 580
 15 <212> DNA
 <213> Homo sapiens

<400> 68
 20 gagtccaggac ggaggacgcg gcaggtcaca gagccacca agtccgaagc tggaagttca 60
 gattctttga tattcaaagg tggatcatct gtgctttttt ttttttatca gtctctcact 120
 ttttatccat catctaattg tgacagctta tttgccttta taccataaga tggggagtag 180
 25 gggtgagatg aaatccaagc atcgtttccc tccccgatg gtcgcctccc tggggtgaga 240
 cggtcgacgt gtcagacttc accaagagca tctccgcct cggtgcagta atgaacttgg 300
 aaacgattta ctccggcact tggttcctgt ctccataaat gcggctgctt taaagggaaat 360
 30 gtaaaaaggg ctgtaaattg gtattgattg ccggtggtct tgaagaaccc caactgagga 420
 ttgaccgttc cttggagtga aggctccgca ttcagacgcc tttcgctta cgtcatacata 480
 35 attgagaagg gaaaggagac gtgttagttt cagtctgatt atttaccatc aaggcataaa 540
 cacttctcag aggcagcggg acccataaaa ccggcccgtg 580

40 <210> 69
 <211> 589
 <212> DNA
 <213> Homo sapiens
 45 <220>
 <221> misc_feature
 <222> ()..()
 <223> "n" refers to an undetermined base

50 <400> 69
 acacgggggg caacctcttg cacctggctc cctgccctcg gtgccacgtt tccagggttc 60
 ctccacgtcg caggctgtgt cagcctcgct ccttccactg cagaattgcy gtccacagcc 120
 55 tggatggggc actctccatg tatccacctg tccctccgtg gctgctgggc tgagtcgctt 180
 ctgatgctaa caagaggcgt ccggctggac taaggccccg gaagctgaga actggagggc 240
 60 aggtgcgggc atcgggcaga gcagctccag caggcaggac ctggggcctc caccctgcac 300

ccctgtgccc cgcgtgtggc ggaaccgccc cgaggggagg ctgtcaccac ggtgacaggc 360
 agccccacgc gagcctgaga accctcagcc cacctttttc tgtaatcaca gcaggcatct 420
 5 ctccggcaag tcaatccagt tccagctggg gctgcctccc ttgcctcatg ggctttatct 480
 tagaactctg agcaataata aaaaagacgc taccgctac aatagatgtg gcagagaatc 540
 10 tggctcttca cttcatcana gatcacctg aaatgatggg tgttggttaa 589

<210> 70
 <211> 748
 <212> DNA
 15 <213> Homo sapiens
 <220>
 <221> misc_feature
 <222> ()..()
 20 <223> "n" refers to an undetermined base

<400> 70
 25 gctacatctn ctctacattc taactaacac ttgttatctt ctgtttttgt ttgtttgttt 60
 ttaatagcca ttctagtagg catgaagtgg tgtttgcctg ctttttttga tggagggtgga 120
 ggaatagggt ggaattgggc ctttaaccatc aattaagctg ggggccttag acctctgtga 180
 30 attggctgtg acaatagcta aaggaggctg ctacctcata ctgaagagat gtttcctaag 240
 tttgtcacgc gagagggcac cgaaccaact tattgtcttg gagggaagaa gcagcaaggc 300
 agaagacttg aacttctcag agaaaaaac agtctacaga cttcatttta tgctgtcctc 360
 35 acacactact gaaagctcta ccttggggac ctggcttgac ttctaacctc cncctgtggt 420
 atttaggaag agctcccagc tgctctgagt ctcagtctcc caatcagtga aatggaggga 480
 40 atagcacctg cctggctgca tcgccccaca gtgctgcaat gagcatccaa cgagagaaag 540
 cttgtcacct gtgttgcaaa ctaagttaca caaatgcagg cagtagcagc tagaagaaaa 600
 tggttgggaa tctgaaaaga attaaagccc cccatgaatt tcttctcacg cctcctccaa 660
 45 aagccaggga ctgcttcacc ccgcctccag gactgctcgc tccagcatct ccggcagctg 720
 ctgacagaat gtatgttgcg gctgtccc 748

50 <210> 71
 <211> 599
 <212> DNA
 55 <213> Homo sapiens
 <220>
 <221> misc_feature
 <222> ()..()
 60 <223> "n" refers to an undetermined base

<400> 71
 gatgactgtt gcccgagctg aggccacgac ccaaccccgga ggaagggaga acagcttccc 60
 atgaagggca tggctgctgc ccataatcc cagggcagga aataaaggga tcttggacta 120
 5 ggcaatcaaa ggacttcctc tccctctaag gccaaaggag aaatgtggct gggactccaa 180
 gctctgtgga tgcctggagg tgccagcagc tggggatcag ctggccccac ctgcagagcc 240
 10 agccagtggg cccctgcat ctccaagggt gggcttatgg gctccaagaa caggtgtttc 300
 tcagggtaac ctgagccctt acaacttcaa ccaagagagt gaaggggagc agccctggag 360
 gccaatgagg agggggatta gtggctactg atgacaaaga catccctgtc ccagagacca 420
 15 gccccttgtg agcagaagaa tggctgccgg gcaaaaggac ctgctatgcc ctccccatac 480
 acatatcatg ncacctgggg accctctgaa taacaggggg cngctttaga gtggcttnat 540
 20 taccaacaag agggccagaa gggctagagc acacgatttc atgntcggcc gcatgncaa 599

<210> 72
 <211> 614
 25 <212> DNA
 <213> Homo sapiens

<400> 72
 gtgcgctatc acgactgttg cccgagctga ggccagaccc aaccccgagg aaggggagaac 60
 30 agcttcccat gaagggcatg gctgctgcca ccataatccc agggcaggaa ataaagggat 120
 cttggactag gcaatcaaag gacttcctct ccctctaagg ccaaggagga aatgtggctg 180
 35 ggactccaag ctctgtggat gcctggaggt gccagcagct ggggatcagc tggccccacc 240
 tgcagagccc agccagtggc tccccctgca tctccaaggt tgggtctatg ggctccaaga 300
 acaggtgttt ctgagggtaa cctcagcccc tacaacttca accaagagag tgaaggggag 360
 40 cagccctgga ggccaatgag gagggggatt agtggctact gatgacaaag acatccctgt 420
 ccccagagcc agccccttgt gagcagaaga atggctgccg gggcaaaagg acctgctatg 480
 45 ccctcccat acacatatca tggcagctgg ggagccctct gaataacagg gggcgcttta 540
 gagtggcttc attaccaaca agaggcccag aaggggctag agccacacga tttcatggtc 600
 ggccgcatgc gcaa 614
 50

<210> 73
 <211> 552
 55 <212> DNA
 <213> Homo sapiens

<400> 73
 aagcggccac agatggccaa gcatgtggag gagagcacia tattttatatt aaatatccaa 60
 60 atacgaacac attcccgcat ggcaccaaca gccgctgaa cagcccgat gccggcttgt 120

gctttttccg ttttgtctag aaatttgggt tgcactaaat tctcagctga atgaagatga 180
gaaggggctg gcagaggggg tggctccagc tctctgagaa cctggctcct tcccgggtgg 240
5 caggagagaga tggcccttgg ggagacgggg aggggtgcact gcctcatgcc caaaccacca 300
gcttctagtt gagaaatcag aattttctct gcagaataag gaaaaagcat tgtcaccatg 360
10 attcacgtgg agctggccac actcaggaaa ttcaatgggg tcccacaggg gctccgaggg 420
ggaaggagag ggcctgggac atgcccctcc agccatcatg gaacaggatg ggaggggccg 480
gccctcactg ctctctaaca gtgaaaagcc acatctccac tttggaaaac acaggcatgt 540
15 gagagcctgg gg 552

<210> 74
<211> 450
20 <212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
25 <222> ()..()
<223> "n" refers to an undetermined base

<400> 74
30 tggaggcttc gagggaagtg aggttccttc ggacacccta gtgggaaggc tccacgcggt 60
aatggaacca cgctgtgaaa cctttgcctt tgggtgtcat ggtggaagca aatcttagaa 120
gacatttaat ttaaaaaatt cagttttaaa aaatgttgac ttaaaaagca gttttgaaaa 180
35 acaacctgga attagcctga gatcgatgcc aactcttagc agtctgtata ctaaacacag 240
ttaaacaact gtagctgctg gcaagctgga acctttttgt aaagaagcac ataaaaagga 300
40 cagaactggg ggaagggtgca ctgggtctttc cacatcgcca ccaggcgttt tgaagcgtgc 360
tgetgacacg ctactcanat gcttctggaa gccaaacaat aanaaaaanc cccattgttt 420
cccttgctgg gttttaccn ccatggtgga 450

<210> 75
<211> 432
50 <212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
55 <222> ()..()
<223> "n" refers to an undetermined base

<400> 75
60 ggacaatgag gagggggtgc acgtggaatc cccacggata ggccggacgc cgggcaggag 60
cctttgcagg ggtgcacagc ctctctgga agccctggtc gctgcctggg gcctgctgca 120

ccctgcgggc tccgcagcgg tggagccagg cctgaactgc ctgctcttgg ccccgctgc 180
 5 ggccctctgc cctttgtctt gcccgtgggg cccggggcct caagctggcc cggggttcct 240
 gaagttagct gacgatgggc tggcctctgg ggctgggtcg tgggccttgt gcaactggccg 300
 ccacgtcacc agcgccaggc ctaccgcgg tgctgctgga gacgcgggat gcccgggctc 360
 10 gggctgtgct ggatcccctg gcgctgcgaa ccccgtagcc ctttccaatc gcgggcnccg 420
 nttaaagccc ga 432

15 <210> 76
 <211> 501
 <212> DNA
 <213> Homo sapiens

20 <220>
 <221> misc_feature
 <222> ()..()
 <223> "n" refers to an undetermined base

25 <400> 76
 gacgagacct agccggcncc atgcgcgcct tgagcctggc gaacagtctg gctggcgcca 60
 cgcgctgat gctcttcgtc cagatcatcc tgatcgacta gaccggcttc catccgagta 120
 30 cgtgctcgct cgatgcgatg tttcgcttgg tggcgaatgg gcaggtagcc ggatcaagcg 180
 tatcgagccg cccgattgca tcagccatga tggatacttt ctcggcagga gcaaggtggg 240
 35 atgacaggag atcctgcccc ggcacttcgc ccaatagcag ccagtccctt cccgcttcag 300
 tgacaacgtc gagcacagct gcccaaggaa cgcccgctgt ggccagccac gatagccgcg 360
 ctgcctcgtc ctgcagttca ttcagggcac cggacaggtc ggtcttgaca aaaagaaccg 420
 40 ggcgcccctg ccgttgacag ccggaacacg gcggcatcag agcagccgat tgtctcgctg 480
 tgcccagtca tagccgaatt c 501

45 <210> 77
 <211> 826
 <212> DNA
 <213> Homo sapiens

50 <400> 77
 gcgcctctgt gggatgacgc accatcctgt ttgtttgcac caagtcattt atctcgtgca 60
 ccccaggggg ccgtgggtccc tgccgggcca tcatgtctgc ttccttatt tgggttttct 120
 55 gccccctcac ttcattttct acttcgcttt tctccttat ccctttgcag tcttgctttt 180
 gggggcattg ctacgccagt aatttgaggg acacctcgtg gagccctagt gtggagccgt 240
 60 cagagcctgg gtaggattct ccgtggtgag gtgctcaggg agacacagga gcattccggc 300

gcctgttcct tgtgcacatc cgcaagtgtc tgcagtgaga ggcatgggtc ccatcttgaa 360
 tgccaacaat gtggcaccca cccccactt gatggggccg agccacagct ggccagggtg 420
 5 accaccatgg acgtgccaga ggcatccgaa acccagctct tgcccagctg ttccactgcc 480
 aactccagcg ttagcaaagc agctctccct tgctttgtct tctacagcag agaacagatt 540
 10 aaaagagaag ctgcaggcag agaaatgcct cttggagcca gatgccccaa aggatctctt 600
 tgaacaaagg gttgctcagg tcagcgtagg ttcttgcat caagcaacaa aatcagagat 660
 gctaacagtt ctcagattca ctccaagtga agactcaaag ctggatttat aaatccccac 720
 15 agagccgctg tgcagaggta gagggccggt ttcaggatga ggaagccctc ttggaagcac 780
 cgtcctccgg ctaacaagcc tccaacctac tgtcggcagg gagaac 826

20 <210> 78
 <211> 433
 <212> DNA
 <213> Homo sapiens

25 <220>
 <221> misc_feature
 <222> ()..()
 <223> "n" refers to an undetermined base

30 <400> 78
 tgcgcagctc cgcgangtgc cggcggggcc cgaccctcag actcgcttgt ccctggagac 60
 caaccctagc gaccaggctc tgccggatcc cgtcgggttt caactcctat tccgaaggctc 120
 35 ctttctcccc taatcacaac acccactcgc ctctttttcc tcctcttcct cagcttccac 180
 cgccgaccgg gcagccccag ttacccgata acggctccca aggccccgtg tttacattct 240
 40 ttccacttgg aagcagaaat tatcacgcc aaattcctac ctgccttccc tggattcctg 300
 gtttcctaag aaacgggttt ggccaccccc tgggcgttcg aacagtccac agaagcgggc 360
 aaaggaaaga cgactcagtc tttccctccc gccaatctct tctccgggac cacagatccc 420
 45 agaagtcacc gcg 433

50 <210> 79
 <211> 424
 <212> DNA
 <213> Homo sapiens

<400> 79
 55 ggcgggcccc accctcagac tcgcttgtcc ctggagacca accctagcga ccaggctctg 60
 ccggatcccc tcgggtttca actcctatcc cgaaggctct ttctccccta atcacaacac 120
 ccaactcgct ctttttcttc ctcttctca gttccaccg ccgaccgggc agccccagtt 180
 60 acccgataac ggctcccaag gccccgtgtt tacattcttt cccactggaa gcagaaatta 240

tcacgccccaa attcctacct gccttccttg gattcctggg ttcctaagaa acggggtttgg 300
 cccacccctg ggcgttcgaa cagtccacag aagcggggcaa aggaaagacg actcagtctt 360
 5 tccccctccgc caatctcttc tccgggacca caaatcccag aagtcaccgc ggccgctaag 420
 ccga 424
 10
 <210> 80
 <211> 285
 <212> DNA
 <213> Homo sapiens
 15
 <220>
 <221> misc_feature
 <222> ()..()
 <223> "n" refers to an undetermined base
 20
 <400> 80
 caaccggggg gcanaggcga tcaaaantgg ggtgcgctgt ggtgggcgac acgtgtggcg 60
 25 cggtgtcat tatccgccct tttcacttcc tggactggaa atggcagacc atatgatggc 120
 aatgaaccac gggcgcttcc ccgacggcac caatgggctg caccatcacc ctgcccaccg 180
 catgggcatg gggcagttcc cgagccccc taccaccag cagcagcagc cccagcacgc 240
 30 cttcaacgcc ctaatgggcg agcacatata ctacggcgcg ggcaa 285
 <210> 81
 35 <211> 401
 <212> DNA
 <213> Homo sapiens
 <400> 81
 40 cagatatgta tctcctcttt tccaaccctg cgtccctttg aggcttggtc ggcgttccca 60
 acctgcccct accccaccaa cccctgtccc tttggccatt agtcccggat tatctagcga 120
 tgccccgtgt accgtctggc tttgctgttt actccgcgct cggccagttg aggctttttg 180
 45 tattttattcc tgattttctc ataggggtaa agtgccttcg ggaggatagg acaagtccca 240
 tctgtttcat acgaattaca gctcggactt cgggcccttt tacactgcct tttgtatctg 300
 50 ttaacttgcg ctaaaaaaga ttcggttctt ttttttgagg aaggggggtg gggggcggag 360
 actctgtcgc ccagtctga gggccgcggc gcgcaagccg a 401
 55 <210> 82
 <211> 268
 <212> DNA
 <213> Homo sapiens
 60 <400> 82
 atagcgcgca caactgtgtc tcttaccag gcacatgcac tatccctgat cccggtgcat 60

gatgggaatg tagtcctgca gccctgtgac caaagggctg ggagtgttta tgagacagca 120
 tctctcagca agcaaagcaa ggctgcaca gccccgcctt ttcctccagt gaggcgcact 180
 5 gttcattaag gagtgttcat gagattacat tttccatcaa gccagccag tcacgcacag 240
 ctctacctct tctctgccc ccccgcaa 268

10 <210> 83
 <211> 989
 <212> DNA
 <213> Homo sapiens

15 <220>
 <221> misc_feature
 <222> ()..()
 <223> "n" refers to an undetermined base

20 <400> 83
 gggtaatggg ggtgaacaga gagggatgcc gaggccagct tgtagtgtgg ctgttggtct 60
 tgtccatcct atggcacaa cctgtcacca cccagatttt gttaggagtc ctcccccaac 120
 25 ttgagagtgg aagctccttt ggcacaaaaa ggggttctgc atcatcccc agccccagc 180
 cctgagcctg ggtctggctc tgaactagac ctccatgaat gaatgcacag catcagtggg 240
 gatccaccat catggggaaa tagtagatac aggaatgatt ttccaaccag attacagact 300
 atttcaagcc cagccagagc ctaccaggcc aacattcccc aggcttgtgc ctctccgagc 360
 ctcagattgc tcatccttca aacgaggagc agctctgctg gcattacctg aactctaggg 420
 30 tcctttataa gctcagactc cagcttagag cacacattga gaggtgctg cccccagag 480
 ccacatacgt gcaacagagg gtggtccaga ccccttattg gtcccatgg ggtttgagag 540
 agaagcctcc agaccagctc aacttctccc tcatctcact taggcctttg cccccagctc 600
 ttaggagggt gtcaggtcac agtgcccat ttcttttctc ttccccagaa atcatgcggg 660
 ggatacctgc tcagacagga ccttcatgaa agccaggctg tgaggtgtgt tggggaatgc 720
 35 ataattgata ggccatcggt cggaggccct cctggaggac caaaatgtaa tcagcagtgg 780
 cgagcttggt cagcagagga attcctttta catcctgggt aggccaaaga cctggcaagc 840
 aagtcctctt ggtcattaaa gaagcatcct gacttgangc aggnacactt aggtcactgc 900
 agccacaaaa atctttgntg ctggattcna aagtaggcat tggggctggg atctgggctc 960
 40 tggcatcctt gancgtgtcg ggggccaaa 989

45 <210> 84
 <211> 250
 <212> DNA
 <213> Homo sapiens

50 <220>
 <221> misc_feature
 <222> ()..()
 <223> "n" refers to an undetermined base

<400> 84
 cgggctcgaa acttcgaaga ccgcggaacc cgaagcngcn cttggctcna atcgcttcgg 60
 5 ctcgagggcgc ccgtncgggt cacgtgaggt gggggcgggc cgaagagggg ggctcccctc 120
 ctctgccgc aggggtggcc gcaagtgcgc ttcaagaggc gcttgatgac ggtaaatgtt 180
 gcagcccga agatgacttt tttctcctcc ttgggttgcg gcaggccgtt agtgggaggt 240
 10 cgcgtcccga 250

<210> 85
 <211> 402
 15 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 20 <222> ()..()
 <223> "n" refers to an undetermined base

<400> 85
 25 ttctcccttg tcattccctt accagagcca cagaaattat ccctgtgggc tcccttgctc 60
 tcactcggcc ttttctggag ttaagagatc caagccaact actgggtctg ttccctgcta 120
 aaatcttagg ccggcgctcc atccacccat ccccatgcct aggactttta agctggcaac 180
 30 ggtacctggg tttagttttc ccttcgtata tcactatctt cgtngcttac cttcttgctc 240
 ctaaagtcc accgatgtgc aaggngatta accactaaag tgcacctgac actactcttg 300
 35 acaaattgca gttgggaggt gagttgatga ctggccggta aatcaaaagt gcttatttag 360
 ggagtgaggg ggcccgcggc anaagccgan ttccagcaca ct 402

<210> 86
 <211> 595
 40 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 45 <222> ()..()
 <223> "n" refers to an undetermined base

<400> 86
 gatcccagaa ggttctggag ccgagtatca gagtttgagc agcgagtcca gcctagcaga 60
 agcgggtgtt gaccggagac ttttcaatgg tgcaaaatga cacactgctt ttgacttggg 120
 55 gatctgtccc ttgtggcacc agaagctaca acaggtncac ctggattcca gctctagctg 180
 gactcggtaa ttgctaagtg ccagctctga agtctgtgat tccgtggaaa tccctttcaa 240
 60 gcccgaaattc tgttttttat gggcctcttg tccaaacagt ttgacttggt aactctgttt 300

ctgtcaagtt gacacttggg cttggcacc cttcatgagc cagatgaaag cggctaaatg 360
 ccogaaaaaa taaaggnttt tacctttttt ttgaaccatt ggtgagcatn taaaaaaatt 420
 5 aggggaaggta aaacccaacc nggncaaacc caactnaaca nttttttttt ccnaaacaag 480
 ggggggctan tttttcactt ggaaaaacaa acaattttta ttgantcttg ananggtgga 540
 naacccaaaat tttttgttgg gttgggttcc gnagnccgaa ttntgcaa at ttctt 595
 10
 <210> 87
 <211> 304
 <212> DNA
 15 <213> Homo sapiens
 <220>
 <221> misc_feature
 <222> ()..()
 20 <223> "n" refers to an undetermined base
 <400> 87
 25 cgtggcccga tgcattcagg gagccctctg tgttggccgc atagcagggtg tagttgccgg 60
 catcctggat gaagacgggc gcgatctgta gacccccga ttcaagaagc atgaacctag 120
 gaatccggac agagccactg gccagaatgt ggttttctaa agaacagtgg agaaaagagg 180
 30 catgttacag tcgtaacgct tgaaggaaat gaagatagtg gttagagcca taagcaagta 240
 atatggttcg gtcctgtgtc cccacccaag tctcgtctng aattgcaatc cccacgtcgg 300
 cgca 304
 35
 <210> 88
 <211> 296
 <212> DNA
 40 <213> Homo sapiens
 <220>
 <221> misc_feature
 <222> ()..()
 45 <223> "n" refers to an undetermined base
 <400> 88
 50 ggctttcgn t aggagttaat ggggcattgg ngggtgggat ggcagggctg ccagcatctg 60
 acccaggagg ctgggaggag gctgctgtgt gaatacacgc tcggcctctc acagtggctg 120
 ccgccgcatt agccccttgt gcttcaggga acagagcatc cgtgatggat gagacttta 180
 55 ttaaagtaat gagacattta taatcgcggt tatctccaaa attaggcctt ttagcaatta 240
 ttcttgggga atattcctcc ggtagatagc tcccttttta gaacaacgtc ggcgca 296
 60
 <210> 89
 <211> 220

<212> DNA
<213> Homo sapiens

<220>
5 <221> misc_feature
<222> ()..()
<223> "n" refers to an undetermined base

10 <400> 89
attggcccgcn caggcgggaa acangctggn nttctctnac cgttntccag cactgcccag 60
accaggagggc gcaggagag gaggggncag cggttccgng accgtcctc ccgtgtccc 120
15 tgctctccag cctntgcctc tgcaggagcc cgcgggantt gcccaggcc cctgtcccca 180
cctgtggctc ccgtcctggt cgctcccggg gccgcggcaa 220

20 <210> 90
<211> 273
<212> DNA
<213> Homo sapiens

25 <220>
<221> misc_feature
<222> ()..()
<223> "n" refers to an undetermined base

30 <400> 90
gnagggnggn ggtgcggac gccgggtggc agttcttgtt cggatgatgtg ggtaaaaaag 60
gactgcagcg aggagccggg gcggcgctcg gagtaatcac cggcggcatc aaaaagcgcc 120
35 atcatggcat cgaggtcgcg gtctgcttgg gagccggtgg cgcgcgcg caaggcagat 180
gcctgcaggc gcataatccag ctcggtagcg ctccatacct cccacaggat ttcttccaca 240
40 gaggcttggg cttgtatagc ctgcgcgcc gca 273

<210> 91
<211> 361
45 <212> DNA
<213> Homo sapiens

<220>
50 <221> misc_feature
<222> ()..()
<223> "n" refers to an undetermined base

<400> 91
55 acggcttctn tnctaagtga cacgggtgtgt gaaattcggg tggggaggta gttctgtaaa 60
ctgcgtctcc ccgccagcta aggaagttga gtgaaggag cgttgccgtc tgggaatcgt 120
agtcctcaca aaggcgtgag taggcggcaa ataaggattt gggtttagcc ttggggattc 180
60 actcctgtca aagctgttag agaagctccc anaactcnta aagtaacaga aactacttgc 240

ggcaacattt gtaacttcca cctgggtcat tatcttccac tgttacctg tggtctagat 300
 aagttataat ttattctaca tatcgttcag aagtcttggt cctgttccat attgtnagca 360
 5 t 361

10 <210> 92
 <211> 462
 <212> DNA
 <213> Homo sapiens

15 <400> 92
 gctgccca cttggatggga aggaccggcg cctgcagcat ctgccctcca agccttcgta 60
 gctccctcct tcttcgagga taaactctaa actccttagc acaacgtggg agccttctca 120
 gagactgggt ccaaccctatc tccagccgca gcctccctc ctggcccccac tgccacaccc 180
 20 ccgggcctcc ggccacactg agcctctccc gggttccag gatacaacac tcgcccattc 240
 atagtgtggt gcctttttgca cgtgctgttc ctctgcttgg ggatgctgtt ggtctttctc 300
 25 agccagggtga agaggacgt gaatgtcacc tgcttgagta tcaggaccgg ggactgggcg 360
 ctggacctag actcttggcc ctggagagaa gccctgcatg gggccgcagc ctgccccgt 420
 ccctgctcac agaaaagctc agccttgcag ccgcgtggga ga 462

30 <210> 93
 <211> 591
 <212> DNA
 35 <213> Homo sapiens

<400> 93
 caaagtcacc tccacggtgc ggctcagcag ctcgccacac ttggtcatgg tgcggggaa 60
 40 ggccgcccctc agctgtaggt gggtagtggc agaacaggag ggtgagggga gactccgaac 120
 tgcctccact tggccgttcc ctccctactg gggggccctg agccagtggc ctctctctc 180
 ggggcctccc cggaaggagc caaggctctgt ctgcgaggca ccggctcccg gccacggcca 240
 45 tcagcccccga gaggtggatc agggcatcac cccactcca cagctgaggc caggggggtca 300
 gggaggcaac cagggcagac ctggaacctg gctctgagac aggacggccg agggcccctc 360
 50 cactctccct ccttcggggt gggcactgac ctggacgcca aagatgtcct cacactggtg 420
 gcgtttgagt agggccact cggacatctg gccctgcagc aggttgggtg agacggccat 480
 ctctccacat gtcacatccg cccgaagcg cttgcagatc cgtcggaagg gcaggttccc 540
 55 aactgcggg gggagcagga cagacacaca tgctcttgca cgcgcacctc a 591

60 <210> 94
 <211> 279
 <212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

5 <222> ()..()

<223> "n" refers to an undetermined base

<400> 94

10 ttntgagttt tggcctgccc acagtctagc cctggacaga gaatccgagg ctgagccatg 60

ctgcagcacc caggacactg catcccagca cctgcccga aatcagccca gggacccaaa 120

15 ggaaagcagg ctccaagctc cccggaagcc aaggaaaata ggaaaacata tcctgccccg 180

gggacacctt ctggaactat gaccacatgc acttgacctt ccggaacaat caccgcatgc 240

acctgacctc ccggaactgt caccaccgcg cgcacctca 279

20

<210> 95

<211> 351

<212> DNA

<213> Homo sapiens

25

<400> 95

cctttattat tggttaaact caccagaaa acccttaact cttagacagc ggctctcatt 60

aagcaaaagg ggaggcacat gaagctccag gcagggccgg gaggggaaccg tgaagccaaa 120

30 ggctctggga gccccaggc acctgcgttt gcattttcat cctggaggag accaggcctc 180

tggtgtgtaa catactcatt 240

35 gattcagtca cctgaccttt gactccatgt attttgttga gtctggatgt gtggtgtgct 300

ctgcccagca gctgggatcc acatgagcac agacatggtc cccccgcggc a 351

40

<210> 96

<211> 171

<212> DNA

<213> Homo sapiens

45

<400> 96

ttgagtgtcg cgtgaatacc taggggacac tcaggggaat gatggctccc ccgagaggta 60

aagggtggaa agaaggggcc tcagcaggtt aggtcttgcg gggctcttct gtagggcgctc 120

50 tgggagatag atccgtgggg ctctagggt cggccctacc cggcgcgggc a 171

<210> 97

<211> 743

55 <212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

60 <222> ()..()

<223> "n" refers to an undetermined base

<400> 97

5 cctccctggc ccttggtccc aaggagcttc ccttggtccc gcctcttcgc cagtgaacttc 60
tcaactggacc attccttttac aaggagcctg ttttttgtgt ttttttttta cacctttttt 120
cttctatttc acagaaggaa caccggacgt ccctntgtga tggcagcagc catgctgcct 180
10 ntgtttccgc tcaggggttc tntgccacct ccaattccac ccagtctntt ggctcggct 240
gggcttcggc tcccgccntt gngccaaaaa ttgcaatgcc cgcggtcagg gcncctttgcg 300
gagtctcacc gcctgaggag gcttgattcc ctctcacag gcagcagcgt ttgatggccg 360
15 gtgacncccc cttttccaag cacatntntc atggcccctg aatgccactt acagggcgtc 420
cctccctgtg ctaagtgtg cctgganctt tgggtgtggc agcagcaaan acctctaccc 480
20 ttgnggatgt tcgtttcggg gnggaaagac anatancaaa gttggtcgta aactgtaaag 540
tgtgctggga ggaaactgag gcagggaggg cctggtgcc ctggggagcn ctgccccgac 600
cccatgtgct tcccaggctc ccttgagacc acgtggatgg cgacttcctg accttgagg 660
25 ccgnggncct cantcctcat gctcgatggc gtcancctcc tcttggggaa atccaancat 720
tcctgacctg aaaatgcacc cnc 743

<210> 98

<211> 589

<212> DNA

<213> Homo sapiens

<400> 98

35 ttgccgcgct gataaaggaa gcgtctagaa ggtctcccca gccttcatca tctgagactt 60
ggctttcagc cccaaagcac taggccctgc tggttaacctt ccaccattaa cttttggtgc 120
40 tcttcaatta gcagcagcca ggggtccttg gcaggatga gaatttgga ggacagcccc 180
agggcatggc ccccggtgc agcaaaagt ctaagtgttc ttctgttgga aggaagccca 240
45 ggagatattg atcagctgca ggtgggggag gcccagatc ccaccttgc ctgcctccag 300
gagaagggttc tccatgggac aaaatggagg cagagtccca ctttgccctgg gcagctccct 360
gagcatggct ccctgtggac ggagctgagt gacgtcatga ctctaggcct caacaaaaga 420
50 gctttggaaa atcccgatga ttcgaattgt attaaatcaa caaacatcg gttgcacagt 480
tactagaaaa cgagatctg cgtcatcact tactagacac gtgacctga acggcggctt 540
55 ccccggtgta aacagcaaag ttctgtaacc cccatgaacg cgctctca 589

<210> 99

<211> 538

<212> DNA

<213> Homo sapiens

<400> 99
 tgccgcgtct gaccctactc tcacaaagac tttccaacta gcataattga gttaaagtgt 60
 5 ccccccaact cccttaattc aagctaaact tgcagtttaa caactatagg agtgatatct 120
 acacattaat gccacacttt aacatgccta acactacaca tgaacacgct tccgggtgct 180
 gttacatccc gctctctccc aagcacgaga cacaggcagg atgctgacgt cctgcttctc 240
 10 tgctgcgggc gggaagtcaa gactccgat ttgctgcagg agttgccgtg gggatcctga 300
 cttcacgcag gagatggtcg gcctctggaa gtgcctggcc cgtttatcct tgaaatctac 360
 15 ctgtgcaggt ggtccttgcc tcagcccctc aggacaacac aggtctttcc taagttacag 420
 ggagaccatc agattgtcgt gtccgagccc cctgaagtgg aaccacagt ctccattcag 480
 tctgccctca gtttccctcc cctctgcagg gccattgctg ctgtggacgc gcctctca 538

<210> 100
 <211> 486
 <212> DNA
 <213> Homo sapiens

<400> 100
 agaggtagaa aaaggagtta gaagcaaaga ggaaaaaata aataaacagg caacaaaaac 60
 30 ccaaccagc cagcctgagc catttgcat agtggtcatt taggaaatta gcagacggga 120
 aacgctgggg agtggagtgg gccccggcct tggggactgc agagcccgct cagccctggg 180
 tggctgggcc cacatgggct gtgccccagg agcacaggag gaccagagg gtggccgaga 240
 35 gagcctcgcc gggctccggt atgggtcctg gcccctcaca ggtgcgagcc tggcccagtg 300
 actgtggacg ctgtgggaga gcaggcctcc gatacgcagg gctgggactg ctgacctgga 360
 40 aggtggtgcc gggcgtgtct ggtgaaggcg ccgttggcag ctagagagag acggcggatg 420
 gggtgacgcc attaccacg gtcccagttt tgaggcttga cggtgacgga aaaggacgtc 480
 ggcgca 486

<210> 101
 <211> 450
 <212> DNA
 <213> Homo sapiens

<400> 101
 aattgaacca ggggtgcacg ccagcgccag acacagtgag cttcatggca actccagttt 60
 55 accggtgaga accatggggc cactcagaga ggcaaagagc ctcacccgag tgagtcctct 120
 ggcttctccc cacctgggcc gggccccagg ccgcgctgtg gttccctttc cagccgtcat 180
 ccctgggtga tgggaggtgg gcattctgtt caaccttgtg ggtcagggag ccagggccag 240
 60 tgtgcagatg agaagaggct gcggttactg gcgatgcgag ggactgtccc cttcgtgggc 300

actttctctt ttgaggccag tgaaatgtgt tccctggggg tgtattcctg agaaggcctc 360

5 atttaaaggg agccgccaaa ccaagtgggc ttagcaaaag cagtttgtca cctggcagca 420

cgtgtgagcc tcgcccggac gcgcctctca 450

10 <210> 102
<211> 292
<212> DNA
<213> Homo sapiens

15 <400> 102
agcgcggcct ggcagattgc ccattaatga aactcagtgg gcagaggctg ctgagggaca 60

cggattccca ctccccgggg gaggggggtgg aaatggcttc ctccctctgc ttccctacca 120

20 ccagtaatgg ggagctcacc atgcttagaa gactcttcct tgcattggagt tcgggcctcc 180

tccctgcacc taccacccta gtggcccca gtcttaaggc tgaaggtaa tcctgtgtcc 240

25 ttcagaagca aaggctgcaa ccgataccaa acagaggctgg ccagcgcggg ca 292

30 <210> 103
<211> 395
<212> DNA
<213> Homo sapiens

35 <220>
<221> misc_feature
<222> ()..()
<223> "n" refers to an undetermined base

40 <400> 103
agagcttatc ccgcgagcac aaggagccg gggcctgggc cgccgtggga aggggtcct 60

45 gccttccggg gacgcggtca gggaagtcca gccggggtgc tctctgcact gcgggtgccg 120

ggctcggcag aggccaacct ggcaaaacga gcaggatctc ccggccccac ctagtgggc 180

50 tccgcctgcc ccaacaacca tcctgccatc ctccctgcga gacaggtgac tttcctctct 240

gatgcggtgc atctgtcatc tgtctaacgg gccattccc cagtgaaca cccccaacca 300

aagacacgaa ggggaaggcg caagcttcta ccaagctcan tttgcccac tggtgcccac 360

ctgcctngta tttggtgact tggaggatag gaagg 395

We claim:

1. A diagnostic or prognostic assay for cancer, comprising:

(a) obtaining a tissue sample from a test tissue;

(b) performing a methylation assay on DNA derived from the tissue sample,

wherein the methylation assay determines the methylation state of a CpG dinucleotide within a DNA sequence of the DNA, and wherein the DNA sequence is a sequence selected from the group consisting of sequences of SEQ ID NOS:1-103, sequences having a nucleotide sequence at least 90% identical to sequences of SEQ ID NOS:1-103, CpG island sequences associated with sequences of SEQ ID NOS:1-103, CpG island sequences associated with sequences having a nucleotide sequence at least 90% identical to sequences of SEQ ID NOS:1-103, and combinations thereof, wherein the CpG island sequence associated with the sequence of the particular SEQ ID NO is that contiguous sequence of genomic DNA that encompasses at least one nucleotide of the particular SEQ ID NO sequence, and satisfies the criteria of having both a frequency of CpG dinucleotides corresponding to an Observed/Expected Ratio >0.6 , and a GC Content >0.5 ; and

(c) determining a diagnosis or prognosis based, at least in part, upon the methylation state of the CpG dinucleotide within the DNA sequence.

2. The diagnostic or prognostic assay of claim 1 wherein the DNA sequence is a sequence selected from the group consisting of CpG island sequences associated with sequences of SEQ ID NOS:1-103, CpG island sequences associated with sequences having a nucleotide sequence at least 90% identical to sequences of SEQ ID NOS:1-103, and combinations thereof.

3. The diagnostic or prognostic assay of claim 2 wherein the DNA sequence is a sequence selected from the group consisting of CpG island sequences associated with sequences of SEQ ID NOS: 2, 4, 6, 7, 9-16, 19, 20, 22-33, 35-43, 48, 51-55, 59, 60, 64, 71, 76, 78-81, 84 and 87-90, and combinations thereof.

4. The diagnostic or prognostic assay of claim 1 wherein the methylation assay procedure is selected from the group consisting of MethyLight, MS-SNuPE, MSP MCA, COBRA, and combinations thereof.

5. The diagnostic or prognostic assay of claim 1 wherein the methylation state of the CpG dinucleotide within the DNA sequence is that of hypermethylation, hypomethylation or normal methylation.

6. The diagnostic or prognostic assay of claim 1 wherein the cancer is selected from the group consisting of bladder cancer, prostate cancer, colon cancer, lung cancer, renal cancer, leukemia, breast cancer, uterine cancer, astrocytoma, glioblastoma, and neuroblastoma.

7. A kit useful for the detection of a methylated CpG-containing nucleic acid comprising a carrier means containing one or more containers comprising:

(a) a container containing a probe or primer which hybridizes to any region of a sequence selected from the group consisting of SEQ ID NOS:1-103, and sequences having a nucleotide sequence at least 90% identical to sequences of SEQ ID NOS:1-103; and

(b) additional standard methylation assay reagents required to affect detection of methylated CpG-containing nucleic acid based, at least in part, on the probe or primer.

8. The kit of claim 7, wherein the additional standard methylation assay reagents are standard reagents for performing a methylation assay from the group consisting of MethyLight, MS-SNuPE, MSP MCA, COBRA, and combinations thereof.

9. The kit of claim 7, wherein the probe or primer comprises at least about 12 to 15 nucleotides of a sequence selected from the group consisting of SEQ ID NOS:1-103, and sequences having a nucleotide sequence at least 90% identical to sequences of SEQ ID NOS:1-103.

10. An isolated nucleic acid molecule comprising a methylated or unmethylated polynucleotide sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:18, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:32, SEQ ID NO:34, SEQ ID NO:37, SEQ ID NO:38, SEQ ID NO:42, SEQ ID NO:44, SEQ ID NO:51, SEQ ID NO:52, SEQ ID NO:62, SEQ ID NO:64, SEQ ID NO:65, SEQ ID NO:68, SEQ ID NO:69, SEQ ID NO:70, SEQ ID NO:71, SEQ ID NO:74, SEQ ID NO:76, SEQ ID NO:82, SEQ ID NO:83, SEQ ID NO:84, SEQ ID NO:86, SEQ ID NO:90, SEQ ID NO:92, SEQ ID NO:97, and SEQ ID NO:100.

11. The nucleic acid of claim 10, wherein the nucleic acid is methylated.

12. The nucleic acid of claim 10, wherein the nucleic acid is unmethylated.

Abstract

There is disclosed 103 novel methylation-altered DNA sequences (“marker sequences”) that have distinct methylation patterns in cancer, compared to normal tissue. In many instances, these marker sequences represent novel sequences not found in the GenBank data base, and none of these marker sequences have previously been characterized with respect to their methylation pattern in human cancers including, but not limited to those of bladder and prostate. These 103 sequences have utility as diagnostic, prognostic and therapeutic markers in the treatment of human cancer, and as reagents in kits for detecting methylated CpG-containing nucleic acids.